



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 179839

TO: Stephen Kapushoc  
Location: REM/3A60/2C70  
Art Unit: 1634  
Wednesday, February 22, 2006  
Case Serial Number: 10/785981

From: Barb O'Bryen  
Location: Biotech-Chem Library  
Remsen 1a69  
Phone: 571-272-2518

*BOB*  
barbara.obryen@uspto.gov

### Search Notes

*swine chip  
requested working folder  
2-23-06*

This Page Blank (uspto)

STIC-Biotech/ChemLib

199839

From: Kapushoc, Stephen T.  
Sent: Thursday, February 16, 2006 4:32 PM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search for 10/785,981

Please search SEQ ID NOs: 1, 2, 3, 4, 5 for application 10/785,981

Thanks,  
Steve

Stephen Kapushoc  
Art Unit 1634 - USPTO  
Tel: 571-272-3312  
Office: REM 3A60  
Mailbox: REM 2C70

RECEIVED  
FEB 16 2006  
(STIC)

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIS: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

This Page Blank (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:36:14 ; Search time 3992.58 Seconds  
(without alignments)  
9396.598 Million cell updates/sec

Title: US-10-785-981-1

Perfect score: 660

Sequence: 1 gagaccagcaatactatgt.....tgaccatgaagcttgagt 660

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	1588	9 BC004067	BC004067 Mus muscu
2	660	100.0	1754	9 BC005660	BC005660 Mus muscu
3	660	100.0	1761	9 BC018485	BC018485 Mus muscu
4	660	100.0	1771	9 BC092053	BC092053 Mus muscu
5	660	100.0	1778	9 BC018223	BC018223 Mus muscu
6	658.4	99.8	1736	9 BC004005	BC004005 Mus muscu
7	658.4	99.8	1756	9 BC092276	BC092276 Mus muscu
8	656.8	99.5	1760	9 BC095965	BC095965 Mus muscu
9	656.8	99.5	1779	9 BC083069	BC083069 Mus muscu
10	653.6	99.0	181014	9 AL840626	AL840626 Mouse DNA
11	647.2	98.1	1722	6 AX306117	AX306117 Sequence
12	647.2	98.1	1722	9 MUSEFTU	M22432 Mus musculu
13	629.6	95.4	145253	9 AC138173	AC138173 Mus muscu
14	629.6	95.4	186564	9 AC119816	AC119816 Mus muscu
15	628.2	95.2	214658	9 AC133509	AC133509 Mus muscu
16	626.4	94.9	1714	9 RNEUFLA	X63561 R.norvegicu
17	626.4	94.9	1737	9 RNEUFLA	X61043 R.norvegicu
18	626.4	94.9	1746	9 BC091297	BC091297 Rattus no

19	626.4	94.9	1771	9 BC063162	BC063162 Rattus no
20	626.4	94.9	1775	9 BC072542	BC072542 Rattus no
21	626.4	94.9	182416	9 AC099724	AC099724 Mus muscu
22	626.4	94.9	192528	9 AC124512	AC124512 Mus muscu
23	626.4	94.9	233219	14 AC136572	AC136572 Rattus no
c 24	626.4	94.2	350880	9 AC102040	AC102040 Mus muscu
25	625.4	94.8	136258	14 AC166149	AC166149 Mus muscu
c 26	625.4	94.8	179579	14 AC166748	AC166748 Mus muscu
27	625.4	94.8	208266	14 AC110913	AC110913 Mus muscu
28	624.8	94.7	1704	9 MMEF1A	X13661 Mouse mRNA
c 29	621.8	94.2	150755	9 AC154274	AC154274 Mus muscu
30	621.6	94.2	1404	9 RAEF1AX	L10339 Rat elongat
31	614.4	93.1	175370	14 AC145861	AC145861 Pan trogl
32	614.4	93.1	175641	9 AC147556	AC147556 Mus muscu
c 33	614.4	93.1	181867	14 AC134666	AC134666 Mus muscu
34	614.4	93.1	22816	14 AC161170	AC161170 Mus muscu
c 35	608.8	92.2	183837	9 AC100406	AC100406 Mus muscu
c 36	608.8	92.2	287411	14 AC154479	AC154479 Mus muscu
37	605.6	91.8	1722	9 CRUEF1A	D00522 Cricetulus
38	596	90.3	1515	4 OCEF1A	X62245 O.cuniculus
39	596	90.3	1718	4 OC09823	U09823 Oryctolagus
c 40	592.6	89.8	242590	14 AC094797	AC094797 Rattus no
41	592.6	89.8	247086	14 AC134118	AC134118 Rattus no
c 42	590.4	89.5	230228	14 AC116224	AC116224 Rattus no
c 43	590	89.4	161222	9 AC154254	AC154254 Mus muscu
c 44	590	89.4	200115	14 AC154774	AC154774 Mus muscu
c 45	588.4	89.2	254169	14 AC127720	AC127720 Rattus no

#### ALIGNMENTS

#### RESULT 1

#### BC004067

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### PUBMED

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

BC004067 1588 bp mRNA linear ROD 07-SEP-2004  
Mus musculus eukaryotic translation elongation factor 1 alpha 1,  
mRNA (CDNA clone MGC:8209 IMAGE:3590965), complete cds.

BC004067

BC004067.1 GI:13278545

MGC.

Mus musculus (house mouse)

Mus musculus

Mus musculus

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muroidae; Muridae; Murinae; Mus.

1 (bases 1 to 1588)

Strausberg R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Schetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullah, S.J., Bosak, S.A., McSwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahey, J., Helton, E., Kettelman, M., Maman, A., Rodriguez, S.,

Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzyzanski, M.I., Skalek, U., Snailus, D.E.,

Schmurch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1588)

Director MGC Project.

Direct Submission

Submitted (28-FEB-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:32:49 ; Search time 560.102 Seconds  
(without alignments)  
7853.391 Million cell updates/sec

Title: US-10-785-981-1

Perfect score: 660

Sequence: 1 gagaccagcaatactatgt.....tgaccatgaagcttgagt 660

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	660	14	ADZ89038 Pig growt
2	660	100.0	660	14	ADZ79331 Swine gro
3	660	100.0	660	14	ADZ77168 Pig growt
4	655.2	99.3	1722	10	ADZ30645 Rat angio
5	647.2	98.1	1722	6	ABI99778 Mouse isc
6	621.6	94.2	1404	10	ADZ58712 Toxicity-
7	621.6	94.2	1404	13	ADV41496 Rat cardi
8	581.6	88.1	1246	4	AAI22127 Probe #12
9	581.6	88.1	1246	4	ABA67206 Human foe
10	581.6	88.1	1246	4	AAI47423 Probe #16
11	581.6	88.1	1246	4	ABA49293 Human bre
12	581.6	88.1	1246	4	ABA34302 Probe #12
13	581.6	88.1	1246	4	AAK41382 Human bon
14	581.6	88.1	1246	4	AAK15648 Human bra
15	581.6	88.1	1246	4	ABS40975 Human liv
16	581.6	88.1	1246	5	AAI07827 Probe #78
17	581.6	88.1	1246	6	ABS15387 Human gen
18	581.6	88.1	1389	12	AD000457 Novel hum
19	581.6	88.1	1389	12	ADN98888 Novel hum

20	581.6	88.1	1494	6	ABL58971	Ab158971 Human tum
21	581.6	88.1	1506	11	ADI31952	Adi31952 Human cDN
22	581.6	88.1	1506	12	ADQ87715	Adq87715 Human tum
23	581.6	88.1	1506	13	ADQ84157	Adq84157 Human tum
24	581.6	88.1	1506	13	ACN38509	Acn38509 Tumour-as
25	581.6	88.1	1506	13	ADZ84019	Adz84019 Human lym
26	581.6	88.1	1696	5	ABA82688	Abas82688 Elongatio
27	581.6	88.1	1696	8	ABX10398	Abx10398 DNA encod
28	581.6	88.1	1696	8	ACC46008	Acc46008 Human elo
29	581.6	88.1	1696	10	ADB98702	Adb98702 Human elo
30	581.6	88.1	1696	10	ADZ82498	Adz82498 Human DNA
31	581.6	88.1	1744	13	ADR07233	Adr07233 Full leng
32	581.6	88.1	1833	8	ACF34510	Acf34510 Gene enco
33	581.6	88.1	1833	12	ADL35468	Adl35468 Human euk
34	581.6	88.1	1833	13	ADS88611	Ads88611 Human hou
35	581.6	88.1	1833	13	ADU60293	Adu60293 Housekeep
36	581.6	88.1	1837	12	ADJ62805	Adj62805 Human cDN
37	581.6	88.1	1837	12	ADF10528	Adf10528 Reference
38	581.6	88.1	1837	14	ADZ26167	Adz26167 Human gen
39	581.6	88.1	1837	14	AEB03626	Aeb03626 Human euk
C 40	581.6	88.1	1950	4	AAI12888	Aai12888 Probe #28
C 41	581.6	88.1	1950	4	ABA54589	Abas4589 Human foe
C 42	581.6	88.1	1950	4	AAI34246	Aai34246 Probe #29
C 43	581.6	88.1	1950	4	ABA44138	Abas44138 Human bre
C 44	581.6	88.1	1950	4	ABA24374	Abas24374 Probe #28
C 45	581.6	88.1	1950	4	AAK28323	Aak28323 Human bon

## ALIGNMENTS

RESULT 1	
ADZ89038	
ID ADZ89038 standard; DNA; 660 BP.	
XX	
AC ADZ89038;	
XX	
DT 28-JUL-2005 (first entry)	
XX	
DE Pig growth factor I (GF I) gene, seq id 1.	
XX	
KW Growth; biochip; swine; hog raising; growth factor I; gene; ds.	
XX	
OS Sus scrofa; Kagoshima Berkshire.	
XX	
PN US2005112597-A1.	
XX	
PD 26-MAY-2005.	
XX	
PF 26-FEB-2004; 2004US-00785981.	
XX	
PR 24-NOV-2003; 2003KR-00083653.	
XX	
PA (KIMC/) KIM C.	
PA (YEOJ/) YEO J.	
PA (LEEJ/) LEE J.	
PA (SONG/) SONG Y.	
PA (CHOK/) CHO K.	
PA (CHUN/) CHUNG K.	
PA (KIMI/) KIM I.	
PA (JINS/) JIN S.	
PA (PARK/) PARK S.	
PA (JUNG/) JUNG J.	
PA (LEEM/) LEE M.	
PA (KWON/) KWON E.	
PA (CHOE/) CHO E.	
PA (CHOH/) CHO H.	
PA (SHIN/) SHIN S.	
PA (NAME/) NAM H.	
PA (HONG/) HONG Y.	
PA (HONG/) HONG S.	
PA (KANG/) KANG Y.	
PA (HAYY/) HA Y.	

```

PA (ROUJ/) ROU J,
PA (KWAC/) KWACK S.
PA (CHOI/) CHOI I.
PA (KIMB/) KIM B.
XX
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;
XX
XX WPI; 2005-403340/41.
XX
XX Functional cDNA chip useful for screening and function analysis of growth
PT specific genes according to breeds and tissues of swine, comprises
PT substrate and probe comprising growth specific genes in muscle and fat
PT tissues of swine.
XX
XX Claim 2; SEQ ID NO 1; 8pp; English.
XX
XX The invention relates to a functional cDNA chip (I) for screening and
CC function analysis of growth specific genes according to breeds and
CC tissues of swine, comprises a probe comprising growth specific genes in
CC muscle and fat tissues of swine, and a substrate on which the probe is
CC immobilized. Further disclosed is a kit (KI) useful for screening and
CC functional analysis of growth specific gene according to breeds and
CC tissues of swine, comprising (I) integrated in it, Cy5-dCTP or Cy3-dCTP
CC bound cDNA from RNA of the tissue to be screened, a fluorescence scanning
CC system, and a computer analysis system. (I) and (KI) are useful for
CC screening and function analysis of growth specific gene according to
CC breeds and tissues of swine. (II) is useful in the swine improvement and
CC breeding of a new breed, and in the hog raising industry. The current
CC sequence represents the pig growth factor I (GF I) gene.
XX
XX Sequence 660 BP; 178 A; 157 C; 166 G; 159 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 660; DB 14; Length 660;
Best Local Similarity 100.0%; Pred. No. 4.2e-193;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGACGAGCAAACTATGTGACCATCATTTGATGCCCGGAGACACAGAGATTTCATCAA 60
DB 1 GAGACGAGCAAACTATGTGACCATCATTTGATGCCCGGAGACACAGAGATTTCATCAA 60
QY 61 AACATGATTACAGGACATCCAGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 AACATGATTACAGGACATCCAGGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 GGTGAATTTGAAGCTGGTATCTCCAGAAACGGGAGAGCCCGGAGCATGCTCTCTGGCT 180
DB 121 GGTGAATTTGAAGCTGGTATCTCCAGAAACGGGAGAGCCCGGAGCATGCTCTCTGGCT 180
QY 181 TACACCTGGGTGTAACACAGCTGATTTGGTGTCAACAAATGGATTCACCGAGCCA 240
DB 181 TACACCTGGGTGTAACACAGCTGATTTGGTGTCAACAAATGGATTCACCGAGCCA 240
QY 241 CCATACAGTCAGAGAGATAGAGGAATCGTTAAGGAAGTCAGCACTCATTAAGAAA 300
DB 241 CCATACAGTCAGAGAGATAGAGGAATCGTTAAGGAAGTCAGCACTCATTAAGAAA 300
QY 301 ATTGGCTACAAACCTTGACACAGTAGCATTTCTGCCAATTTCTGGTTGGATGGTGAACAAC 360
DB 301 ATTGGCTACAAACCTTGACACAGTAGCATTTCTGCCAATTTCTGGTTGGATGGTGAACAAC 360
QY 361 ATGCTGGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAAGTCACCGCAAGAT 420
DB 361 ATGCTGGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAAGTCACCGCAAGAT 420
QY 421 GGCAGTGCAGTGCAGCACCAGCTGTAAGCTTGGATTTGATCTTATCCACCACTTCGT 480
DB 421 GGCAGTGCAGTGCAGCACCAGCTGTAAGCTTGGATTTGATCTTATCCACCACTTCGT 480
QY 481 CCAACTGACAAGCCTCTCGGACTGCCCTCCAGGATGTCTATAAAATGGAGGCAATGGC 540
DB 481 CCAACTGACAAGCCTCTCGGACTGCCCTCCAGGATGTCTATAAAATGGAGGCAATGGC 540

```

---

```

QY 541 ACTGTCCTCTGGGCGGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
DB 541 ACTGTCCTCTGGGCGGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
QY 601 GCTCCAGTCAATGTAACTCAACTGAAGTCAAGTCTGTGTAATGCAACCATGAAGCTTTGAGT 660
DB 601 GCTCCAGTCAATGTAACTCAACTGAAGTCAAGTCTGTGTAATGCAACCATGAAGCTTTGAGT 660
RESULT 2
ADZ79331
ID ADZ79331 standard; cDNA; 660 BP.
XX
AC ADZ79331;
XX
XX 28-JUL-2005 (first entry)
XX
XX Swine growth factor GE-I cDNA for cDNA chip diagnostic method.
XX
XX biochip; screening; EST; expressed sequence tag; probe; muscle;
KW immobilization; gene expression; polymorphism; diagnosis;
KW animal breeding; growth factor; ss.
XX
XX Unidentified.
XX
XX US2005112602-A1.
XX
XX 26-MAY-2005.
XX
XX 27-FEB-2004; 2004US-00789723.
XX
XX 24-NOV-2003; 2003KR-00083651.
XX
XX (KIMC/) KIM C.
PA (YEOJ/) YEO J.
PA (LEEJ/) LEE J.
PA (SONG/) SONG Y.
PA (CHOK/) CHO K.
PA (CHUN/) CHUNG K.
PA (KIMI/) KIM I.
PA (JINS/) JIN S.
PA (PARK/) PARK S.
PA (JUNG/) JUNG J.
PA (LEEM/) LEE M.
PA (KWON/) KWON E.
PA (CHOE/) CHO E.
PA (CHOH/) CHO H.
PA (SHIN/) SHIN S.
PA (NAMH/) NAM H.
PA (HONG/) HONG Y.
PA (HONG/) HONG S.
PA (KANG/) KANG Y.
PA (HAYY/) HA Y.
PA (ROUJ/) ROU J.
PA (KWAC/) KWACK S.
PA (CHOI/) CHOI I.
PA (KIMB/) KIM B.
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;
XX
XX WPI; 2005-371656/38.
XX
XX New cDNA chip comprising a probe capable of detecting marker genes
PT specifically expressed in the muscle and fat tissues of swine, and a
PT substrate, useful for screening and function analysis of swine genes.
XX
XX Claim 9; SEQ ID NO 1; 15pp; English.
XX
XX The invention relates to a cDNA chip for screening and function analysis
CC of swine genes comprising a probe capable of detecting marker genes

```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:51:20 ; Search time 3910.41 Seconds  
(without alignments)  
7896.735 Million cell updates/sec

Title: US-10-785-981-1

Perfect score: 660

Sequence: 1 gagaccagcaatactatgt.....tgccaccatgaagcttgagt 660

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hic:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gss1:\*  
10: gb\_gss2:\*  
11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	660	100.0	1700	4	BC092292 Mus muscu
2	656.8	99.5	673	3	BM945443 UI-M-EMO-
3	656.8	99.5	678	8	CK788670 HESC3_101
4	656.8	99.5	710	6	CB248368 UI-M-FD0-
5	656.8	99.5	719	3	BM945776 UI-M-EMO-
6	656.8	99.5	746	3	BM946759 UI-M-EMO-
7	656.8	99.5	748	3	BQ042422 UI-M-EMO-
8	656.8	99.5	749	6	CA322009 UI-M-EMO-
9	656.8	99.5	755	5	BQ444687 UI-M-EMO-
10	656.8	99.5	757	7	CF951596 UI-M-HL0-
11	656.8	99.5	760	7	CF949906 UI-M-HK0-
12	656.8	99.5	764	3	BQ177933 UI-M-EMO-
13	656.8	99.5	771	3	BQ177724 UI-M-EMO-
14	656.8	99.5	808	5	BQ571059 UI-M-PB0-
15	656.8	99.5	879	5	BUS12595 AGENCOURT
16	656.8	99.5	926	6	CA979421 AGENCOURT
17	656.8	99.5	928	5	BQ922583 AGENCOURT
18	656.8	99.5	1111	3	BM453687 AGENCOURT
19	656.8	99.5	1145	3	BM461263 AGENCOURT
20	656.8	99.5	1153	5	BUS15235 AGENCOURT
21	656.8	99.5	1733	4	AK081725 Mus muscu
22	656.8	99.5	1734	4	AK076696 Mus muscu

23	656.8	99.5	1736	4	AK083361 Mus muscu
24	655.8	99.4	690	7	CN456673 UI-M-HP0-
25	655.8	99.4	701	7	CK780521 UI-M-RAO-
26	655.8	99.4	708	6	CD364228 UI-M-GL0-
27	655.8	99.4	748	6	CF734957 UI-M-RAO-
28	655.2	99.3	686	7	CN455811 UI-M-HP0-
29	655.2	99.3	693	7	CF951163 UI-M-HL0-
30	655.2	99.3	879	6	CA978866 AGENCOURT
31	655.2	99.3	931	6	CB183269 AGENCOURT
32	655.2	99.3	1722	4	BC003969 Mus muscu
33	655.2	99.3	1731	4	AK032914 UI-M-EMO-
34	654.8	99.2	745	3	BM946164 UI-M-EMO-
35	654.2	99.1	749	6	CD364147 UI-M-GL0-
36	654.2	99.1	749	3	CN456691 UI-M-HP0-
37	654.2	99.1	943	5	BQ891120 AGENCOURT
38	653.8	99.1	760	6	CF731605 UI-M-RAO-
39	653.8	99.1	776	7	CK780620 UI-M-EMO-
40	653.6	99.0	781	3	BM945492 UI-M-EMO-
41	653.2	99.0	725	6	CF731619 UI-M-HL0-
42	653.2	99.0	803	7	CF950931 UI-M-RAO-
43	652	98.8	910	6	CA978558 AGENCOURT
44	651.8	98.8	745	2	BE911061 601662158
45	651.6	98.7	795	5	BQ571020 UI-M-PB0-

#### ALIGNMENTS

RESULT 1	BC092292	1700 bp	musculus	linear	HTC 04-APR-2005
LOCUS	Mus musculus cdna clone IMAGE:5068840, containing frame-shift errors.				
DEFINITION	BC092292.1 GI:62185631				
ACCESSION	BC092292.1				
VERSION	GI:62185631				
KEYWORDS	HTC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 1700)				
AUTHORS	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Hsieh, F., Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Loquellano, N.A., Peters, G.J., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
CONSTRM	Mammalian Gene Collection Program Team				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1700)				
AUTHORS	NIH MGC Project				
CONSTRM	Direct Submission				
TITLE	Submitted (01-APR-2005) National Institutes of Health, Mammalian				
JOURNAL	Gene Collection (MGC), Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
COMMENT	Contact: MGC help desk				

Email: cgapbs-rc@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc\_mcg@nigri.nih.gov  
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
 Torgerson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 193 Row: n Column: 8  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, GenomeScan gene prediction, Similarity but not identity  
 to protein

This clone has the following problem: frame shifted.

#### FEATURES

source  
 1. 1700  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech 11"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5068840"  
 /tissue\_type="Mammary tumor metastasized to lung. Tumor  
 arose spontaneously from a senescent normal mammary  
 (clonal) outgrowth infected with the virus MMTV."  
 /clone\_lib="NCI CGAP\_Lu29"  
 /lab\_hosts="DH10B"  
 /note="vector: pcMV-SPORT6"

#### ORIGIN

Query Match 100.0%; Score 660; DB 4; Length 1700;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-185;  
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GAGACGACAAATCTATGTGACCATCATGATGCTCCCGGACGACACAGAGACTTCATCAAA 60
    |||
Db 253 GAGACGACAAATCTATGTGACCATCATGATGCTCCCGGACGACACAGAGACTTCATCAAA 312
    |||
Qy 61 AACATGATTACAGGCACATCCAGGCTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
    |||
Db 313 AACATGATTACAGGCACATCCAGGCTGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
    |||
Qy 121 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGACCCGCGAGCATGCTCTCTGGCT 180
    |||
Db 373 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGACCCGCGAGCATGCTCTCTGGCT 432
    |||
Qy 181 TACACCTGGGTGAAACAGCTGATTGTTGGTGTCAACAAAATGGATTTCCACGAGCCA 240
    |||
Db 433 TACACCTGGGTGAAACAGCTGATTGTTGGTGTCAACAAAATGGATTTCCACGAGCCA 492
    |||
Qy 241 CCATACAGTCAGAGAGATACGAGAAATCGTTAAGGAAGTCAGACCTACATTAAGAAA 300
    |||
Db 493 CCATACAGTCAGAGAGATACGAGAAATCGTTAAGGAAGTCAGACCTACATTAAGAAA 552
    |||
Qy 301 ATTGGCTACACCCCTGCACAGTAGTACATTTGTGCCAATTTCTGTTGGAATGGTGAACAC 360
    |||
Db 553 ATTGGCTACACCCCTGCACAGTAGTACATTTGTGCCAATTTCTGTTGGAATGGTGAACAC 612
    |||
Qy 361 ATGCTGAGGCCAAGTGTCTAATATGCTTGGTTCAAGGGATGGAAGTCACCCGAAAGAT 420
    |||
Db 613 ATGCTGAGGCCAAGTGTCTAATATGCTTGGTTCAAGGGATGGAAGTCACCCGAAAGAT 672

```

```

Qy 421 GGCAGTGCAGTGGCACCACCGCTGCTGGAGCTTTGGATTGTATCTTACCACCAACTCGT 480
    |||
Db 673 GGCAGTGCAGTGGCACCACCGCTGCTGGAGCTTTGGATTGTATCTTACCACCAACTCGT 732
    |||
Qy 481 CCAACTGACAAAGCCTCTGCGACTGCCCCCTCCAGGATGTCTATAAAATTTGGAGGCAATGGC 540
    |||
Db 733 CCAACTGACAAAGCCTCTGCGACTGCCCCCTCCAGGATGTCTATAAAATTTGGAGGCAATGGC 792
    |||
Qy 541 ACTGTCCCTGTGGCCGAGTGGAGACTGTGTTCTCAAACTGGCATGGGTACCTTTT 600
    |||
Db 793 ACTGTCCCTGTGGCCGAGTGGAGACTGTGTTCTCAAACTGGCATGGGTACCTTTT 852
    |||
Qy 601 GCTCCAGTCAATGTAACAACCTGAAGTCAAGTCTGTTGAAATGCACCATGAAGCTTTGAGT 660
    |||
Db 853 GCTCCAGTCAATGTAACAACCTGAAGTCAAGTCTGTTGAAATGCACCATGAAGCTTTGAGT 912
    |||

```

#### RESULT 2

BM945443  
 LOCUS  
 DEFINITION  
 UI-M-EMO-bvh-1-15-0-UI.r1 NIH\_BMAP\_EMO Mus musculus cDNA clone  
 IMAGE:5692190 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS

SOURCE  
 ORGANISM

Mus musculus (house mouse)  
 Mus musculus  
 Mus musculus  
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS

TITLE  
 COMMENT

NIH-MGC <http://mgs.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
 Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

Seq primer: pyX-5.

Location/Qualifiers  
 1. 673  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5692190"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 18.5 dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_EMO"

/note="Organ: brain; Vector: pyX-Asc; Site\_1: EcoR I;  
 Site\_2: Not I; The library was constructed according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured mRNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was size selected according to mRNA size fraction,  
 ligated with EcoR I adaptor, digested with Not I, and then  
 cloned directionally into pyX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA  
 tail, is CAGCCAGCAG. This library was created for the  
 University of Iowa Mouse Brain Molecular Anatomy Project  
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous  
 System', supported by National Institute of Mental Health  
 (NIMH), Hemin Chin, Ph.D., program coordinator."

#### ORIGIN

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C	1	581.6	88.1	1506	3	US-09-023-655-1378 Sequence 1278, App
	2	581.6	88.1	1506	3	US-09-949-016-204 Sequence 204, App
	3	581.6	88.1	1747	3	US-09-949-016-1833 Sequence 1833, App
	4	581.6	88.1	1749	3	US-09-949-016-3699 Sequence 3699, App
	5	581.6	88.1	1749	3	US-09-949-016-3700 Sequence 3700, App
	6	581.6	88.1	5749	3	US-09-949-016-15441 Sequence 15441, A
	7	581.6	88.1	5749	3	US-09-949-016-15442 Sequence 15442, A
	8	580	87.9	1753	9	5225348-2 Patent No. 5225348
	9	570.4	86.4	2128	3	US-08-371-377-16 Sequence 16, Appl
	10	570.4	86.4	2128	3	US-08-875-553D-22 Sequence 22, Appl
C	11	562.4	85.2	1750	3	US-09-949-016-2972 Sequence 2972, App
	12	562.4	85.2	1750	3	US-09-949-016-2973 Sequence 2973, App
	13	562.4	85.2	5750	3	US-09-949-016-14714 Sequence 14714, A
	14	562.4	85.2	5750	3	US-09-949-016-14715 Sequence 14715, A
	15	458.4	69.5	546	3	US-09-385-982-466 Sequence 466, App
	16	423.4	64.2	1229	3	US-09-023-655-1368 Sequence 1368, App
	17	420.6	63.7	1747	3	US-09-949-016-15448 Sequence 1548, App
	18	401.6	60.8	533	3	US-09-328-111-469 Sequence 469, App
	19	371.4	56.3	601	3	US-09-949-016-22182 Sequence 22182, A
	20	371.4	56.3	601	3	US-09-949-016-131604 Sequence 131604, A
C	21	371.4	56.3	601	3	US-09-949-016-131608 Sequence 131608, A
	22	363	55.0	85963	3	US-09-949-016-13804 Sequence 13804, A
	23	340.6	51.6	1377	3	US-09-614-221A-492 Sequence 492, App
	24	339	51.4	1500	3	US-09-174-768-3 Sequence 3, Appli



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:17:34 ; Search time 798.947 Seconds  
(without alignments)  
6831.225 Million cell updates/8

**Title:** US-10-785-981-1

**Perfect score:**

Sequence: 1 gagaccagcaatactatgt.....tgcaccatgaagcctttgagt 660

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 BEQS, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0% Maximum Match 100%

Maximum price 100%  
Listing first 45 summaries

**Database :**

```
Published_Applications NA Main:
1: /cgn2_6/ptodata/1/pubnpa/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubnpa/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubnpa/US09_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubnpa/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubnpa/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubnpa/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubnpa/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubnpa/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubnpa/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubnpa/US11_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query				ID	Description
	Score	Match	Length	DB		
1	660	100.0	660	9	US-10-785-981-1	Sequence 1, Appli
2	660	100.0	660	9	US-10-789-723-1	Sequence 1, Appli
3	660	100.0	660	9	US-10-786-052-1	Sequence 1, Appli
4	657.4	99.6	1040	7	US-10-424-599-16011	Sequence 16011, A
5	655.2	99.3	1722	6	US-10-316-253-208	Sequence 208, App
6	581.6	88.1	1246	3	US-09-864-761-19622	Sequence 19622, A
7	581.6	88.1	1494	8	US-10-416-330-11	Sequence 11, Appl
8	581.6	88.1	1506	7	US-10-641-643-1278	Sequence 1278, Ap
9	581.6	88.1	1696	6	US-10-374-979-74	Sequence 74, Appl
10	581.6	88.1	1696	7	US-10-182-936A-74	Sequence 74, Appl
11	581.6	88.1	1696	8	US-10-477-238A-653	Sequence 653, App
12	581.6	88.1	1696	8	US-10-680-287A-653	Sequence 653, App
13	581.6	88.1	1696	9	US-10-477-173-653	Sequence 653, App
14	581.6	88.1	1720	8	US-10-278-698-242	Sequence 242, App
15	581.6	88.1	1720	8	US-10-278-698-756	Sequence 756, App
16	581.6	88.1	1833	7	US-10-231-956A-76	Sequence 76, Appl
17	581.6	88.1	1833	8	US-10-684-422-254	Sequence 254, App
18	581.6	88.1	1833	9	US-10-489-740-65	Sequence 65, Appl
19	581.6	88.1	1837	7	US-10-955-157-623	Sequence 623, App
20	581.6	88.1	1837	7	US-10-439-703-75	Sequence 75, Appl
21	581.6	88.1	1950	3	US-09-864-761-2840	Sequence 2840, Ap
22	581.6	88.1	3415	3	US-09-823-245A-438	Sequence 438, App
23	578.4	87.6	959	9	US-10-617-316-38	Sequence 38, Appl

## ALIGNMENTS

## RESULT 1

US-10-785-981-1

Sequence 1, Application US/10785981

; Publication No. US20050112597A1

GENERAL INFORMATION:

APPLICANT: GYEONGSANGNAM-DO

APPLICANT: Chulwook, KIM

**; TITLE OF INVENTION: SCREENING EXPRESSION**

10; TITLE OF INVENTION: AND FUNCTIONAL CDNA CHIP PREPARED BY USING THE SAME

FILE REFERENCE: 3884-0120P

; CURRENT APPLICATION NUMBER: US/10/785,981

;  
; CURRENT FILING DATE: 2004-06

; PRIOR APPLICATION NUMBER:

PRIOR FILING DATE: 2003-11-24

; NUMBER OF :

; SOFTWARE: pa  
CEO ID NO 1

```

; SEQ ID NO 1
: 1 ENCTLY: 550

```

LENGTH: 660  
TYPE: DNA

TYPE: DNA

```

Query Match      100.0%; Score 660; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 5e-208;
Matches 660: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

```
Qy 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Db 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Qy 361 ATGCTGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAGTCAACCCGCAAGAT 420
Db 361 ATGCTGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAGTCAACCCGCAAGAT 420
Qy 421 GGCAAGTGCAGTGGCACCACGCTGCTGGAAGCTTTGGATTGTATCTTACCACCACTCGT 480
Db 421 GGCAAGTGCAGTGGCACCACGCTGCTGGAAGCTTTGGATTGTATCTTACCACCACTCGT 480
Qy 481 CCAACTGACAAGCTCTGCGACTGCCCTCCAGATGCTCTATAAAATTGGAGGCAATGGC 540
Db 481 CCAACTGACAAGCTCTGCGACTGCCCTCCAGATGCTCTATAAAATTGGAGGCAATGGC 540
Qy 541 ACTGTCCCTGTGGCCGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
Db 541 ACTGTCCCTGTGGCCGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
Qy 601 GCTCCAGTCAATGTAACTGAAGTCAAGTCTGTTGAAATGCAACCAATGAAGCTTTGAGT 660
Db 601 GCTCCAGTCAATGTAACTGAAGTCAAGTCTGTTGAAATGCAACCAATGAAGCTTTGAGT 660

RESULT 2
US-10-789-723-1
; Sequence 1, Application US/10789723
; Publication No. US20050112602A1
; GENERAL INFORMATION:
; APPLICANT: GYEONGSANGNAM-DO
; APPLICANT: Chulwook, KIM
; TITLE OF INVENTION: cDNA chip for screening specific genes and analyzing their
; FILE OF INVENTION: function in swine
; FILE REFERENCE: YLOP040109US
; CURRENT APPLICATION NUMBER: US/10789,723
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: KR 2003-83651
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Kagoshima Berkshire
US-10-789-723-1

Query Match 100.0%; Score 660; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 5e-208;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGACCAGCAAACTACTATGTGACCATCATTTGATGCCCGGACACAGAGACTTCATCAAA 60
Db 1 GAGACCAGCAAACTACTATGTGACCATCATTTGATGCCCGGACACAGAGACTTCATCAAA 60
Qy 61 AACATGATTACAGGCACATCCAGGCTGACTGTCTGCTGATTTGCTGCTGGTGT 120
Db 61 AACATGATTACAGGCACATCCAGGCTGACTGTCTGCTGATTTGCTGCTGGTGT 120
Qy 121 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGCCCGGACATGCTCTTCTGGCT 180
Db 121 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGCCCGGACATGCTCTTCTGGCT 180
Qy 181 TACACCTGGGTGTGAACAGCTGATTGTTGGTGTCAACAAAATGGATTCCACCGAGCCA 240
Db 181 TACACCTGGGTGTGAACAGCTGATTGTTGGTGTCAACAAAATGGATTCCACCGAGCCA 240
Qy 241 CCATACAGTCAGAGAGATACAGGAAATCGTTAAGAACTCAGCACTCATTAAGAAA 300
Db 241 CCATACAGTCAGAGAGATACAGGAAATCGTTAAGAACTCAGCACTCATTAAGAAA 300
Qy 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Db 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
```

```
Db 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Qy 361 ATGCTGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAGTCAACCCGCAAGAT 420
Db 361 ATGCTGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAGTCAACCCGCAAGAT 420
Qy 421 GGCAAGTGCAGTGGCACCACGCTGCTGGAAGCTTTGGATTGTATCTTACCACCACTCGT 480
Db 421 GGCAAGTGCAGTGGCACCACGCTGCTGGAAGCTTTGGATTGTATCTTACCACCACTCGT 480
Qy 481 CCAACTGACAAGCTCTGCGACTGCCCTCCAGATGCTCTATAAAATTGGAGGCAATGGC 540
Db 481 CCAACTGACAAGCTCTGCGACTGCCCTCCAGATGCTCTATAAAATTGGAGGCAATGGC 540
Qy 541 ACTGTCCCTGTGGCCGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
Db 541 ACTGTCCCTGTGGCCGAGTGGAGACTGGTGTCTCAAACTGGCATGGTGGTTACCTTT 600
Qy 601 GCTCCAGTCAATGTAACTGAAGTCAAGTCTGTTGAAATGCAACCAATGAAGCTTTGAGT 660
Db 601 GCTCCAGTCAATGTAACTGAAGTCAAGTCTGTTGAAATGCAACCAATGAAGCTTTGAGT 660

RESULT 3
US-10-786-052-1
; Sequence 1, Application US/10786052
; Publication No. US20050113568A1
; GENERAL INFORMATION:
; APPLICANT: GYEONGSANGNAM-DO
; APPLICANT: Chulwook, KIM
; TITLE OF INVENTION: NOVEL GROWTH RELATED GENES FROM SWINE
; FILE REFERENCE: 3884-0119P
; CURRENT APPLICATION NUMBER: US/10786,052
; CURRENT FILING DATE: 2004-02-26
; PRIOR APPLICATION NUMBER: KR 2003-83652
; PRIOR FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Kagoshima Berkshire
US-10-786-052-1

Query Match 100.0%; Score 660; DB 9; Length 660;
Best Local Similarity 100.0%; Pred. No. 5e-208;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGACCAGCAAACTACTATGTGACCATCATTTGATGCCCGGACACAGAGACTTCATCAAA 60
Db 1 GAGACCAGCAAACTACTATGTGACCATCATTTGATGCCCGGACACAGAGACTTCATCAAA 60
Qy 61 AACATGATTACAGGCACATCCAGGCTGACTGTCTGCTGATTTGCTGCTGGTGT 120
Db 61 AACATGATTACAGGCACATCCAGGCTGACTGTCTGCTGATTTGCTGCTGGTGT 120
Qy 121 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGCCCGGACATGCTCTTCTGGCT 180
Db 121 GGTGAATTTGAAGCTGGTATCTCAAGAACGGGACAGCCCGGACATGCTCTTCTGGCT 180
Qy 181 TACACCTGGGTGTGAACAGCTGATTGTTGGTGTCAACAAAATGGATTCCACCGAGCCA 240
Db 181 TACACCTGGGTGTGAACAGCTGATTGTTGGTGTCAACAAAATGGATTCCACCGAGCCA 240
Qy 241 CCATACAGTCAGAGAGATACAGGAAATCGTTAAGAACTCAGCACTCATTAAGAAA 300
Db 241 CCATACAGTCAGAGAGATACAGGAAATCGTTAAGAACTCAGCACTCATTAAGAAA 300
Qy 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Db 301 ATTGGCTACAACTGACAGTAGCAATTTGTCCTGTTGGAATGTTGACAAC 360
Qy 361 ATGCTGAGGCAAGTGTCTAATATGCTTGGTTTCAAGGGATGGAAGTCAACCCGCAAGAT 420
```

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

mis Page Blank (uspto)







Qy	241	CCATACAGT	CAGAGAGAT	CGAGAAAT	CGTTAAGGAAGT	CAGCACCT	CTACATTTAAGAAA	300
Db	481	CCCTACAGC	CAAGAGAT	TCAGGAAAT	TGTTAAGGAAGT	CAGCACCT	TTCATTTAAGAAA	540
Qy	301	ATTGGCTAC	ACCCCTG	CACAGT	AGCATTTGTC	CCAAATTTCTGGTT	TGGAAATGGTCACAC	360
Db	541	ATTGGCTAC	AAACCCCG	CACAGT	AGCATTTGTC	CCAAATTTCTGGTT	TGGAAATGGTCACAC	600
Qy	361	ATGCTGGAG	CCAAGT	GTCTAAT	TGCTTGGTT	CAAGGGAT	GGAAAGTCA	420
Db	601	ATGCTGGAG	CCAAGT	GTCTAAT	TGCTTGGTT	CAAGGGAT	GGAAAGTCA	660
Qy	421	GGCAGTGC	AGTGGCA	CCACGCT	GTGTGAAG	CTTTGGAT	TGTATCCT	480
Db	661	GGCAATGC	AGTGGAA	CCACGCT	GTGTGAAG	CTTTGGAT	TGTATCCT	720
Qy	481	CCAACTGAC	AAAGCCT	CTCGGAC	TGCCCTCC	CAGGATGCT	TATAAATTT	540
Db	721	CCAACTGAC	AAAGCCT	CTCGGAC	TGCCCTCC	CAGGATGCT	TATAAATTT	780
Qy	541	ACTGTCCT	GTGGGCG	AGTGGAG	ACTGGTGT	CTCAAACT	TGGCATG	600
Db	781	ACTGTCCT	GTGGGCG	AGTGGAG	ACTGGTGT	CTCAAACT	TGGCATG	840
Qy	601	GCTCCAGT	CAATGTA	CAACTG	GAAGTCA	AGTCTG	TTGAAATG	660
Db	841	GCTCCAGT	CAATGTA	CAACTG	GAAGTCA	AGTCTG	TTGAAATG	900

## RESULT 2

US-10-947-249-133  
; Sequence 133, Application US/10947249  
; Publication No. US20050287541A1  
; GENERAL INFORMATION:  
; APPLICANT: Akira NAKAGAWARA  
; APPLICANT: Miki OHIRA  
; APPLICANT: Shin ISHII  
; APPLICANT: Takeeshi GOTO  
; APPLICANT: Hiroyuki KUBO  
; APPLICANT: Takahiro HIRATA  
; APPLICANT: Yasuko YOSHIDA  
; APPLICANT: Saichi YAMADA  
; TITLE OF INVENTION: Microarray for Predicting the Prognosis of Neuroblastoma  
; FILE REFERENCE: 117007  
; CURRENT APPLICATION NUMBER: US/10/947,249  
; CURRENT FILING DATE: 2004-09-23  
; PRIOR APPLICATION NUMBER: US 60/505,614  
; PRIOR APPLICATION NUMBER: 2003-09-25  
; NUMBER OF SEQ ID NOS: 200  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 133  
; LENGTH: 1837  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-947-249-133

Qy	181	TACACCTGGGTGGAACAGCTGATTTGGTGTCAACAAAATGGATTCACCCGAGCCA	240
Db	483	TACACACTGGGTGGAACCAACTAAITTTGTCGGTGTAAACAAAATGGATTCACCTGAGCCA	542
Qy	241	CCATACAGTCAGAGAGATACGAGGAATCGTTAAGGAAGTCAGCACCTACATTAAAGAAA	300
Db	543	CCCTACAGCCAGAAGATATGAGGAATTCGTTAAGGAAGTCAGCACTTACATTAAAGAAA	602
Qy	301	ATTGGCTACAACCCCTGCACAGTAGCAATTTCTGGTTCACAAATTTCTGGTTGGAATGGTGACAAC	360
Db	603	ATTGGCTACAACCCCGACACAGTAGCAATTTCTGGTTCACAAATTTCTGGTTGGAATGGTGACAAC	662
Qy	361	ATGCTGAGGCCAAGTGTCTAATATATGCTTGGTTCAAGGGATGGAAGTCACCCGCAAAAGAT	420
Db	663	ATGCTGAGGCCAAGTGTCTAATATGCTTGGTTCAAGGGATGGAAGTCACCCGTTAAGGAT	722
Qy	421	GGCAGTGGCAGTGGCAACACGCTCTGGAAAGCTTTGGATTGTATCTCTACCAACCAACTCGT	480
Db	723	GGCAATGGCAGTGGAAACCAAGCTCTGGAGCTCTGGACTGCATCCTACCAACCAACTCGT	782
Qy	481	CCAACCTGACAAAGCCTCTGCGACTGCCCCCTCCAGGATGCTATATAAATTTGGAGGCAATTGGC	540
Db	783	CCAACCTGACAAAGCCTCTGCGCTCTCCAGGATGCTATATAAATTTGGTGTATTTGGT	842
Qy	541	ACTGTCCCTGTGGCCGAGTGGAGACTGGTGTCTCAAACCTGGCATGGTGGTTACCTTT	600
Db	843	ACTGTCTCTGTTGGCCGAGTGGAGACTGGTGTCTCAAACCCGGTATGGTGGTCACTTT	902
Qy	601	GCTCCAGTCAATGTAAACACTGAAGTCGAAGTCTGTTGAAATGCACCATGAAGCTTTGAGT	660
Db	903	GCTCCAGTCAACGTTTACAAACGAAGTAAATCTGCGAAATGCACCATGAAGCTTTGAGT	962
RESULT 3			
US-11-000-688-278			
; Sequence 278, Application US/11000688			
; Publication No. US20050287544A1			
; GENERAL INFORMATION:			
; APPLICANT: BERTUCCI, Francois			
; APPLICANT: HOULGATTE, Remi			
; APPLICANT: BIRNBAUM, Daniel			
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS			
; FILE REFERENCE: 1423-R-03			
; CURRENT APPLICATION NUMBER: US/11/000,688			
; CURRENT FILING DATE: 2004-12-01			
; PRIOR APPLICATION NUMBER: US 60/525,987			
; PRIOR FILING DATE: 2003-12-01			
; NUMBER OF SEQ ID NOS: 1596			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 278			
; LENGTH: 1837			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial sequences:primer			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)..(1837)			
; OTHER INFORMATION: eukaryotic translation elongation factor 1			
; OTHER INFORMATION: alpha 1(BEF1a1) gene.			
US-11-000-688-278			

### RESULT 3

```

US-11-000-688-278
; Sequence 278, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAY
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 278
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1837)
; OTHER INFORMATION: eukaryotic translation elongation factor 1
; OTHER INFORMATION: alpha 1(EEF1A1) gene.
US-11-000-688-278

```

	Query Match	88.1%;	Score 581.6;	DB 12;	Length 1837;
	Best Local Similarity	92.6%;	Pred. No. 9.1e-166;		
	Matches 611;	Conservative 0;	Mismatches 49;	Indels 0;	Gaps 0;
Qy	1	GAGACCGCAAATACTATGTGACCATCATTTGATGCCCGGACACACAGAGACTTCATCAAA	60		
Db	303	GAGACCGCAAGTACTATGTGACTATCATTTGATGCCCGGACACACAGAGACTTTTATCAAA	362		
Qy	61	AACATGATTACAGGCACATCCAGGCTGACTGTCTCTCTGATTTGTCGCTGGTGTT	120		
Db	363	AACATGATTACAGGCACATCTCAGGCTGACTGTCTCTCTCTGATTTGTCGCTGGTGTT	422		

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:36:14 ; Search time 3206.16 Seconds  
(without alignments)  
9396.598 Million cell updates/sec

Title: US-10-785-981-2

Perfect score: 530

Sequence: 1 gctgactgacggagaatc.....tagtgcaattgaaatcctgg 530

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5881141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	530	100.0	5929	4	AB025261	- AB025261 Sus scrofa
2	526.8	99.4	5866	4	AB025262	--AB025262 Sus scrof
3	482	90.9	5987	4	AB059399	AB059399 Bos tauru
4	477.2	90.0	5982	4	AB088366	AB088366 Equus cab
5	470.8	88.8	1953	8	BC022376	BC022376 Homo sapi
6	470.8	88.8	5925	6	CS018093	CS018093 Sequence
7	470.8	88.8	5925	6	CS025608	CS025608 Sequence
8	470.8	88.8	5925	6	CS119311	CS119311 Sequence
9	470.8	88.8	5925	8	AF111785	AF111785 Homo sapi
10	470.8	88.8	5926	6	CQ731010	CQ731010 Sequence
11	459.6	86.7	5988	4	OCU32574	U32574 Oryctolagus
12	450	84.9	5932	4	AB025260	AB025260 Sus scrof
13	431.2	81.4	4910	9	AK220548	AK220548 Mus muscu
14	424.4	80.1	1806	5	CR523587	CR523587 Gallus ga
15	424.4	80.1	5904	6	CQ733076	CQ733076 Sequence
16	424.4	80.1	6016	6	CS018091	CS018091 Sequence
17	424.4	80.1	6016	6	CS025606	CS025606 Sequence
18	424.4	80.1	6016	8	AF111783	AF111783 Homo sapi

19	424.4	80.1	31111	5	CHRMVHE	J02714 Chicken emb
20	421.6	79.5	6010	8	HSMYOSIN	Z38133 H.sapiens m
21	421.6	79.5	6016	6	CS071067	CS071067 Sequence
22	421.6	79.5	6016	8	HUMMYOHP	M36769 Homo sapien
23	421.6	79.5	6018	6	CQ731011	CQ731011 Sequence
24	421.2	79.5	5595	9	BC008538	BC008538 Mus muscu
25	421.2	79.5	5965	4	AB059398	AB059398 Bos tauru
26	419.6	79.2	5951	5	AB021180	AB021180 Gallus ga
27	416.4	78.6	2409	8	BC093082	BC093082 Homo sapi
28	416.4	78.6	5956	6	CS018089	CS018089 Sequence
29	416.4	78.6	5956	8	AF111784	AF111784 Homo sapi
30	416.4	78.6	5957	6	CQ729334	CQ729334 Sequence
31	416.4	78.6	6085	6	CS025604	CS025604 Sequence
32	413.2	78.0	6010	5	AF272034	AF272034 Gallus ga
33	410	77.4	6031	5	AY116217	AY116217 Gallus ga
34	408.8	77.1	5980	4	AB088365	AB088365 Equus cab
35	406.8	76.8	5874	5	GU87231	GU87231 Gallus gall
36	403.6	76.2	6022	5	AY116218	AY116218 Gallus ga
37	395.6	74.6	6067	5	AF272033	AF272033 Gallus ga
38	366.8	69.2	2520	5	AF240689	AF240689 Rana pipi
39	365.2	68.9	2520	5	AF240690	AF240690 Rana pipi
40	362.4	68.4	6032	6	AX770500	AX770500 Sequence
41	362.4	68.4	6032	8	HSMYHC	X13988 Human mRNA
42	357.2	67.4	633	9	AF336977	AF336977 Mus muscu
43	354	66.8	3333	5	BC044194	BC044194 Danio rer
44	352.4	66.5	5913	5	AY921650	AY921650 Danio rer
45	348.6	65.8	6005	5	BC076678	BC076678 Xenopus t

#### ALIGNMENTS

RESULT 1  
AB025261  
LOCUS AB025261 5929 bp mRNA linear MAM 26-JAN-2001  
DEFINITION Sus scrofa mRNA for myosin heavy chain 2b, complete cds.  
ACCESSION AB025261  
VERSION AB025261.1 GI:5360747  
KEYWORDS myosin heavy chain 2b.  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa

REFERENCE 1 (sites)  
AUTHORS Chikuni,K., Tanabe,R., Muroya,S. and Nakajima,I.  
TITLE Differences in molecular structure among the porcine myosin heavy chain-2a, -2x, and -2b isoforms  
JOURNAL Meat Sci 57, 311-317 (2001)  
REFERENCE 2 (bases 1 to 5929)  
AUTHORS Chikuni,K.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-1999) Koichi Chikuni, National Institute of Animal Industry, Meat Science Laboratory; Nourin-kenkyu-danchi,P.O.box 5, Tsukuba, Ibaraki 305-0901, Japan (E-mail:chikuni@naii.affrc.go.jp, Tel:81-298-38-8686, Fax:81-298-38-8606)

FEATURES  
source Location/Qualifiers  
1..5929  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/strain="Landrace"  
/db\_xref="taxon:9823"  
/sex="male"  
/tissue\_type="skeletal muscle"  
/dev\_stage="adult"  
54..5867  
/codon\_start=1  
/product="myosin heavy chain 2b"  
/protein\_id="BAA82145.1"  
/db\_xref="GI:5360748"  
CDS  
translation="MSSDQEMAIFGEAAPYLKSEKERIEAQNKPFDAKTSVFVAEPK  
ESFVKGTVQSGREGKVTVKTEAGATLTVKEDQVFPNPPKFDKIEDAMMTHLEPAV

LYNLKERYAAMIYTYSGLFVTVNPKYKWLVPVNAEVVTVAYRGKQAEAPPHIFSISD  
NAVOPLMTDRENOSILITGSGAGKTVNTKRVIOYFATIAVTGCKKEBTPCKMOGT  
LEQOLIISANPLIEAFGNKTVNRDNRSGFKGFIIRHPTGTGLASADIEYLLKESRV  
TFQLKABRSHTIIFYQIMNSKKPELIELMLITTPDYAFVFSQSEITVPSIDDOELMA  
TDSAIETLGTSDERSVSIYKLTGAHVHYNLKFQKQREBPQAPDGTGEVADKAAYIQG  
LNSADLLKALCYPRVKGNEFVTKGTVQOVYNAGALAKAVYDRKFLMWVTRINQOL  
DTKPORYFTGLVDIAGFBI FDNPSLEQLCINFTEKLOQFNFHMFVLQOEYKKEG  
IEWEFDGMDLAACTELIBKPMGIFSIIEECWFPKATDTSFNKLYEOHLGKSNF  
OKPKPAKGAEAHPSLIHYAGTVDNVITGWLKDNKOPLANETVVLGYOKSVKTLAFL  
AERQSEEGTGKGGKGGSGFQVTSALFRENKLMNLRSTHPIHVRCLIIENETK  
PGAMEHELVLHQRCNGVLEGIRICRKGFPISILYADFQORYKVYNASAIPEGOFIDS  
KXASCKLLGSDIDHTQYKFGHTKVFVFKAGLGTLEEMRDEKLAQLITRTQACRGLF  
MYREFRQWERRRESIFCIQYINRAFNMVHPWPKLYFKIKPLKSAETEKEMANNKE  
EFKTKEDLASAKRKEBEKEMVMOEKNDLQLOVQAEADGLADEBCRDOLIKTK  
IOLEAKTKEVTERADEEENAEITAKKRKLEDECSSELKWDIDDLLETLAKVEKHA  
TENKVNLTSEMAGLDENIAKLITKEKALQAEHQOITLDLOAEDKVNLTITKATKLE  
QOVDDLEGSLEQKRLMDLERAKRLEGDLKLAQESTMDIENDKQDLDEKLLKGF  
MSNLQKIEDEALAMLOKKIKELQARTLEEBEETEAERASRAKAEKORSLSRLE  
EISERLEPAGGATSAQIEMNKKRAEFQKMRDLERATLQHEATAALRKKHADSVAE  
LGSQIDNLQVOKLKEKSELKMEIDDLASMETVSKAKGNLEKMCRTLEDOLSEVK  
TKEEHOQLINELSAOKARLOTBSEGFESQOLDEKALVSQLSRGKQATQOIELEKRO  
LEBETKASALAHAVQSRHDCDLLEHQYEEBQAEKALQRAMSKANSEVQWTKYE  
TDAIQRTLEEELAAQLQDAEBEHVAVNAKASLEKTKORLQNEVEDMLDVER  
SNAACALDKQNFDFKILAEWKHYEETQAELEASQKESRSSTELFKVKNAYEBSL  
DOLETLKRENKIQEISDLTEQIAEGGKHIHELEKVKQIEQKSELQALAEABAS  
LEHEEGKILRIOLELNOVKSEIDRKIAEKDEEIDOMKRNHIRVVEGSMQTLDAEIRSR  
NDALRIKKMEGDLNEMEIQLNHNRAQATEAIRNLNTQGVLDQTLDDAIRGDD  
LKQGLAVVERRANLQAEIBELRASLEQTERSRRAVEQELLDASERVOLLHTQNTSLI  
NTRKKLETDISQIQEGMEDIVQEARNAEKAKAITDAAMAELEKKEQDTSAHLERM  
KXMEQTVLQHLRDLAEQALGKGGKKQIKLEARVRELENEVEKQENKRAVGL  
RKHERRVKELTYOTEDRKNVRLQDLVDLQSKVKAYKQAEABEQSNVLSKFRK  
LQHELEBAERADIAESQVKNLKVKSREVHTKVISE"

## ORIGIN

Query Match 100.0%; Score 530; DB 4; Length 5929;  
Best Local Similarity 100.0%; Pred. No. 1.7e-135;  
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCAGGAGATCCGGGCGAGAAAGAC	60
Db	551	GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCAGGAGATCCGGGCGAGAAAGAC	610
Qy	61	TGTGAACACGAAAGCGTGTCTATCCAGTACTTTGGCCACAAATCGCCGTCACTGGGAGAGAA	120
Db	611	TGTGAACACGAAAGCGTGTCTATCCAGTACTTTGGCCACAAATCGCCGTCACTGGGAGAGAA	670
Qy	121	GAAGGAGAAACCTACTCTCTGCAAAATGACGGGACCTCTGGAAGATCAGATCATCATGTC	180
Db	671	GAAGGAGAAACCTACTCTCTGCAAAATGACGGGACCTCTGGAAGATCAGATCATCATGTC	730
Qy	181	CAACCCCTGCTCGAGGCTTTGGCAAGCGCCAGACCGTGAGGACGACCACTCTCTCG	240
Db	731	CAACCCCTGCTCGAGGCTTTGGCAAGCGCCAGACCGTGAGGACGACCACTCTCTCG	790
Qy	241	CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTGGGAAGCTGGCTTCTGCTGACAT	300
Db	791	CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTGGGAAGCTGGCTTCTGCTGACAT	850
Qy	301	CGAAACATATCTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGACAGAAAGACTA	360
Db	851	CGAAACATATCTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGACAGAAAGACTA	910
Qy	361	CCACATTTTTTATCAGATCATGTCTAAACAGAACCCAGAGCTCATTTGAAATGCTCCTGAT	420
Db	911	CCACATTTTTTATCAGATCATGTCTAAACAGAACCCAGAGCTCATTTGAAATGCTCCTGAT	970
Qy	421	CACCACCAACCATATGACTACGCTTCGTGAGTCAAGGGAGAGTCACTGTGCCCGACAT	480
Db	971	CACCACCAACCATATGACTACGCTTCGTGAGTCAAGGGAGAGTCACTGTGCCCGACAT	1030
Qy	481	TGATGACCAAGAGAGCTGTAGGCGCACAGATAGTGCCATTTGAAATCTCTGG	530
Db	1031	TGATGACCAAGAGAGCTGTAGGCGCACAGATAGTGCCATTTGAAATCTCTGG	1080

## RESULT 2

AB025262  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AB025262  
Sus scrofa mRNA for myosin heavy chain 2x, complete cds.  
AB025262  
AB025262.1 GI:5360749  
myosin heavy chain 2x.  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;  
Sus.

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

## FEATURES

## source

## CDS

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:32:49 ; Search time 449.779 Seconds  
(without alignments)  
7853.391 Million cell updates/sec

Title: US-10-785-981-2

Perfect score: 530

Sequence: 1 gctgactgacgggagaaac.....tagtgccattgaaatcctgg 530

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	530	100.0	530	14	ADZ89039 Pig growt
2	530	100.0	530	14	ADZ79332 Swine gro
3	530	100.0	530	14	ADZ77169 Pig growt
4	470.8	88.8	5925	10	ADZ29577 Human tum
5	470.8	88.8	5925	10	ADW95930 cDNA enco
6	470.8	88.8	5925	14	ADX15796 DNA enco
7	470.8	88.8	5925	14	ADX25983 Novel cel
8	432.4	81.6	5874	5	AAS85420 DNA enco
9	424.4	80.1	6016	14	ADW95928 cDNA enco
10	424.4	80.1	6016	14	ADX15794 DNA enco
11	424.4	80.1	6016	14	ADX25985 Novel cel
12	421.6	79.5	6010	12	ADQ22038 Human sof
13	421.6	79.5	6010	12	ADQ22007 Human sof
14	421.6	79.5	6010	12	ADQ17241 Human sof
15	421.6	79.5	6010	12	ADQ17288 Human sof
16	421.6	79.5	6016	4	AAS7390 Human sof
17	421.2	79.5	5595	14	ADX26120 Novel cel
18	416.4	78.6	5956	14	ADW95926 cDNA enco
19	416.4	78.6	6085	14	ADX15792 DNA enco

20	362.4	68.4	6032	8	ACA89885	Aca89885 Gene diff
21	352.8	66.6	6035	14	ADX26230	Adx26230 Novel cel
22	347.6	65.6	5992	13	ADQ80266	Adq80266 Skeletal
23	328.4	62.0	2477	14	AEB50291	Aeb50291 Human myo
24	328.4	62.0	5395	14	AEB50288	Aeb50288 Human myo
25	328.4	62.0	5661	2	AAV21518	Aav21518 Rattus no
26	328.4	62.0	5930	10	ADB52833	Adb52833 Primary r
27	328.4	62.0	5930	13	ADV41112	Adv41112 Rat cardi
28	328.4	62.0	5930	14	ADX26221	Adx26221 Novel cel
29	328.4	62.0	6057	14	AEB50285	Aeb50285 Human myo
30	328.4	62.0	6177	14	AEB50286	Aeb50286 Human myo
31	328.4	62.0	6290	14	AEB50284	Aeb50284 Human myo
32	328.4	62.0	6903	14	AEB50292	Aeb50292 Human myo
33	328.4	62.0	8921	14	AEB50283	Aeb50283 Human myo
34	328.4	62.0	9119	14	AEB50287	Aeb50287 Human myo
35	320.4	60.5	5886	13	ADS97707	Ads97707 Rabbit al
36	320.4	60.5	12801	13	ADS97709	Ads97709 Rabbit al
37	319.2	60.2	6054	14	ADX26110	Adx26110 Novel cel
38	313.2	59.1	5918	11	ACN92725	Acn92725 Breast ca
39	308	58.1	4775	10	ADK18363	Adk18363 Human NOV
40	308	58.1	5780	10	ADK18365	Adk18365 Human NOV
41	308	58.1	6008	10	ADK18367	Adk18367 Human NOV
42	308	58.1	6008	12	ADP13461	Adp13461 Renal cel
43	308	58.1	6008	14	ADX25975	Adx25975 Novel cel
44	301.6	56.9	5925	12	ADP72935	Adp72935 Renal tox
45	301.6	56.9	5925	13	ADV41113	Adv41113 Rat cardi

## ALIGNMENTS

RESULT 1

ADZ89039

ID ADZ89039 standard; DNA; 530 BP.

XX

AC ADZ89039;

XX

DT 28-JUL-2005 (first entry)

XX

DE Pig growth factor II (GF II) gene, seq id 2.

XX

KW Growth; biochip; swine; hog raising; growth factor II; gene; ds.

XX

OS Sus scrofa; Kagoshima Berkshire.

XX

PN US2005112597-A1.

XX

PD 26-MAY-2005.

XX

PF 26-FEB-2004; 2004US-00785981.

XX

PR 24-NOV-2003; 2003KR-00083653.

XX

PA (KIMC/) KIM C.

PA (YEOJ/) YEO J.

PA (LEEJ/) LEE J.

PA (SONG/) SONG Y.

PA (CHOK/) CHO K.

PA (CHUN/) CHUNG K.

PA (KIMI/) KIM I.

PA (JINS/) JIN S.

PA (PARK/) PARK S.

PA (JUNG/) JUNG J.

PA (LEEM/) LEE M.

PA (KWON/) KWON E.

PA (CHOE/) CHO E.

PA (CHOH/) CHO H.

PA (SHIN/) SHIN S.

PA (NAMEH/) NAM H.

PA (HONG/) HONG Y.

PA (HONG/) HONG S.

PA (KANG/) KANG Y.

PA (HAYY/) HA Y.

```
PA (ROUJ/) ROU J.
PA (KWAC/) KWACK S.
PA (CHOI/) CHOI I.
PA (KIMB/) KIM B.
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;
XX WPI: 2005-403340/41.
XX Functional cDNA chip useful for screening and function analysis of growth
XX specific genes according to breeds and tissues of swine, comprises
PT substrate and probe comprising growth specific genes in muscle and fat
PT tissues of swine.
XX
XX Claim 2; SEQ ID NO 2; 8pp; English.
XX
XX The invention relates to a functional cDNA chip (I) for screening and
XX function analysis of growth specific genes according to breeds and
XX tissues of swine, comprises a probe comprising growth specific genes in
XX muscle and fat tissues of swine, and a substrate on which the probe is
XX immobilized. Further disclosed is a kit (KI) useful for screening and
XX functional analysis of growth specific gene according to breeds and
XX tissues of swine, comprising (I) integrated in it, Cy5-dCTP or Cy3-dCTP
XX bound cDNA from RNA of the tissue to be screened, a fluorescence scanning
XX system, and a computer analysis system. (I) and (KI) are useful for
XX screening and function analysis of growth specific gene according to
XX breeds and tissues of swine. (I) is useful in the swine improvement and
XX breeding of a new breed, and in the hog raising industry. The current
XX sequence represents the pig growth factor II (GF II) gene.
XX
XX Sequence 530 BP; 154 A; 137 C; 127 G; 112 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 530; DB 14; Length 530;
Best Local Similarity 100.0%; Pred. No. 7e-159;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTGACTGATCGGAGGAATCAGTCTATCTTAATCACCGGAGATCCGGGGCAGGAAGAC 60
DB 1 GCTGACTGATCGGAGGAATCAGTCTATCTTAATCACCGGAGATCCGGGGCAGGAAGAC 60
QY 61 TGTGAACACGAAGCGTGTCTCAGTACTTTGGCCCAATCGCGTCACTGGGGAGAAGAA 120
DB 61 TGTGAACACGAAGCGTGTCTCAGTACTTTGGCCCAATCGCGTCACTGGGGAGAAGAA 120
QY 121 GAAGGAGGAACCTACTCTCTGCAAAATGCAGGGGACTCTGGAAGATCAGATCATCAGTGC 180
DB 121 GAAGGAGGAACCTACTCTCTGCAAAATGCAGGGGACTCTGGAAGATCAGATCATCAGTGC 180
QY 181 CAACCCCTGCTCAGGCGCTTTGGCCAAAGCGGCAAGACCGTGTAGGAAACGACAACTCTCTCG 240
DB 181 CAACCCCTGCTCAGGCGCTTTGGCCAAAGCGGCAAGACCGTGTAGGAAACGACAACTCTCTCG 240
QY 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTGGGAAGCTGCTCTGCTGACAT 300
DB 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTGGGAAGCTGCTCTGCTGACAT 300
QY 301 CGAAACATATCTCTAGAGAAGTCTAGAGTCACTTTCCAGCTAAAGGCGAGAAAGACTA 360
DB 301 CGAAACATATCTCTAGAGAAGTCTAGAGTCACTTTCCAGCTAAAGGCGAGAAAGACTA 360
QY 361 CCACATTTTATCAGATCATGTCTTAACAAGACCCAGAGCTCATTTGAAATGCTCTGAT 420
DB 361 CCACATTTTATCAGATCATGTCTTAACAAGACCCAGAGCTCATTTGAAATGCTCTGAT 420
QY 421 CACCACCAACCCATATCAGTACGCTTCGTCAGTCAAGGGGAGATCACTGTCGCCAGCAT 480
DB 421 CACCACCAACCCATATCAGTACGCTTCGTCAGTCAAGGGGAGATCACTGTCGCCAGCAT 480
QY 481 TGATGACCAAGAGGAGCTGTGGCCACAGATAGTGCCATTGAAATCTCTGG 530
DB 481 TGATGACCAAGAGGAGCTGTGGCCACAGATAGTGCCATTGAAATCTCTGG 530
```

RESULT 2  
ADZ79332  
ID ADZ79332 standard; cDNA; 530 BP.  
XX  
AC ADZ79332;  
XX  
DT 28-JUL-2005 (first entry)  
DE  
XX Swine growth factor GB-II cDNA for cDNA chip diagnostic method.  
XX biochip; screening; EST; expressed sequence tag; probe; muscle;  
KW immobilization; gene expression; polymorphism; diagnosis;  
KW animal breeding; growth factor; ss.  
XX  
OS Unidentified.  
XX  
PN US2005112602-A1.  
XX  
PD 26-MAY-2005.  
XX  
PF 27-FEB-2004; 2004US-00789723.  
XX  
PR 24-NOV-2003; 2003KR-00083651.  
XX  
XX (KIMC/) KIM C.  
PA (YEOJ/) YEO J.  
PA (LEEJ/) LEE J.  
PA (SONG/) SONG Y.  
PA (CHOK/) CHO K.  
PA (CHUN/) CHUNG K.  
PA (KIMI/) KIM I.  
PA (JINS/) JIN S.  
PA (PARK/) PARK S.  
PA (JUNG/) JUNG J.  
PA (LEEM/) LEE M.  
PA (KWON/) KWON E.  
PA (CHOE/) CHO E.  
PA (CHOH/) CHO H.  
PA (SHIN/) SHIN S.  
PA (NAMH/) NAM H.  
PA (HONG/) HONG Y.  
PA (HONG/) HONG S.  
PA (KANG/) KANG Y.  
PA (HAYI/) HA Y.  
PA (ROUJ/) ROU J.  
PA (KWAC/) KWACK S.  
PA (CHOI/) CHOI I.  
PA (KIMB/) KIM B.  
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;  
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;  
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;  
XX  
DR WPI: 2005-371656/38.  
XX  
XX New cDNA chip comprising a probe capable of detecting marker genes  
PT specifically expressed in the muscle and fat tissues of swine, and a  
PT substrate, useful for screening and function analysis of swine genes.  
XX  
XX Claim 9; SEQ ID NO 2; 15pp; English.  
XX  
XX The invention relates to a cDNA chip for screening and function analysis  
XX of swine genes comprising a probe capable of detecting marker genes  
XX specifically expressed in the muscle and fat tissues of swine and a  
XX substrate on which the probe is immobilized. The invention also includes  
XX a kit for screening and function analysis of swine genes comprising the  
XX cDNA chip. The cDNA chip is useful for screening and function analysis of  
XX swine genes, particularly for comparing genetic expression according to  
XX swine breeds and tissues, genetic mutation screening, genetic  
XX polymorphism interpretation, development of new drugs for disease  
XX treatment and disease diagnosis, swine improvement. This sequence

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:51:20 ; Search time 3140.18 Seconds  
(without alignments)  
7896.735 Million cell updates/sec

Title: US-10-785-981-2

Perfect score: 530

Sequence: 1 gctgactgacgggagatc.....tagtgccattggaatctcgg 530

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	462.8	87.3	5667	11	DQ052205 Homo sapi
2	462.8	87.3	5667	11	DQ052206 Homo sapi
3	450	84.9	2674	4	AK041122 Mus muscu
4	447.6	84.5	752	8	AK023235 LIB4216-0
5	431.2	81.4	3339	4	AK029328 Mus muscu
6	429.2	81.0	623	8	AK023628 LIB4216-0
7	427.6	80.7	648	8	DN421412 LIB4216-0
8	427.6	80.7	771	8	DN425844 LIB4216-0
9	424.4	80.1	5820	11	DQ029772 Homo sapi
10	422.4	79.7	650	8	DN424241 LIB4216-1
11	421.6	79.5	5814	11	DQ052207 Homo sapi
12	418	78.9	5814	11	DQ052208 Homo sapi
13	416.4	78.6	6062	4	BSX10904 Homo sapi
14	415.2	78.3	785	8	DN423007 LIB4216-0
15	412.4	77.8	675	6	CD621454 560287790
16	398.8	75.2	710	8	DN421704 LIB4216-0
17	374.8	70.7	817	8	CK901213 JGI CAAM9
18	369	69.6	493	5	BU387960 603858632
19	368.2	69.5	5805	11	DQ029773 Pan trogl
20	367.8	69.4	539	1	AW918573 EST349877
21	367	69.2	649	8	DN421699 LIB4216-0
22	366.6	69.2	667	8	DN421346 LIB4216-0

23	364.6	68.8	659	3	BM486201
24	364	68.7	763	8	DN426850
25	362.4	68.4	5823	11	DQ036335
26	354.4	66.9	558	7	CN698824
27	354.4	66.9	626	7	CN699265
28	351.6	66.3	591	7	CR756940
29	350.8	66.2	933	7	CK017852
30	346	65.3	847	8	DN935535
31	344.6	65.0	5823	11	DQ036336
32	343.2	64.8	5802	11	DQ029771
33	338.2	63.8	599	8	DN423594
34	336.8	63.5	839	8	CX390733
35	336.4	63.5	5826	11	DQ051102
36	335	63.2	442	1	AV599364
37	335	63.2	719	7	CK687877
38	334.8	63.2	551	3	BM179887
39	331	62.5	774	6	CF287168
40	329.4	62.2	395	5	BU946705
41	328.8	62.0	864	8	CX328746
42	328.4	62.0	643	8	DN376020
43	324.6	61.2	818	8	CX312040
44	324.2	61.2	5802	11	DQ029770
45	322	60.8	603	6	CD499600

#### ALIGNMENTS

RESULT 1  
DQ052205  
LOCUS Homo sapiens MYH1 gene, VIRTUAL TRANSCRIPT, partial sequence, 5667 bp DNA linear GSS 02-JUN-2005  
DEFINITION Homo sapiens MYH1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.  
ACCESSION DQ052205  
VERSION DQ052205.1 GI:66905676  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 5667)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.  
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees  
JOURNAL (er) Plos Biol. 3 (6), E170 (2005)  
PUBMED 15869325  
REFERENCE 2 (bases 1 to 5667)  
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B., Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.  
FEATURES  
Location/Qualifiers  
source 1..5667  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
gene <1..>5667  
/gene="MYH1"  
ORIGIN /locus\_tag="HC20221"

Query Match 87.3%; Score 462.8; DB 11; Length 5667;  
Best Local Similarity 92.9%; Pred. No. 1.3e-125;  
Matches 485; Conservative 0; Mismatches 37; Indels 0; Gaps 0;





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:15:29 ; Search time 130.047 Seconds  
(without alignments)  
7244.353 Million cell updates/sec

Title: US-10-785-981-2  
Perfect score: 530  
Sequence: 1 gctgactgacgggagaatc.....tagtgccattgaaatcctgg 530

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PPTRUS\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PPTRUS\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470.8	88.8	5925	3	US-09-949-016-5233
2	424.4	80.1	6016	3	US-09-949-016-1054
3	424.4	80.1	6016	3	US-09-949-016-2264
4	362.4	68.4	6011	3	US-09-949-016-3017
5	347.6	65.6	5992	3	US-09-949-016-546
6	347.6	65.6	5992	3	US-09-949-016-2263
7	328.4	62.0	5661	3	US-08-938-105-2
8	308	58.1	6008	3	US-09-949-016-5058
9	144	27.2	6644	3	US-08-875-435B-5
10	141.4	26.7	5919	3	US-08-875-435B-2
11	141.2	26.6	6175	3	US-08-875-435B-1
12	141	26.6	5883	3	US-09-949-016-5001
13	139.2	26.3	5574	3	US-09-917-254-40
14	139.2	26.3	6861	3	US-09-949-016-1240
15	139.2	26.3	6861	3	US-09-949-016-1241
16	139.2	26.3	6861	3	US-09-949-016-1242
17	123.8	23.4	7596	3	US-09-023-655-1463
18	122.2	23.1	7453	3	US-09-949-016-4965
19	109.4	20.6	3581	3	US-09-949-016-16975
20	105.2	19.8	28355	3	US-09-949-016-16975
21	105	19.8	4688	3	US-09-949-016-4522
22	98.2	18.5	7501	3	US-09-620-312D-249
23	94.6	17.8	30271	3	US-09-949-016-12796
24	94.6	17.8	30272	3	US-09-949-016-14006

ALIGNMENTS

RESULT 1

US-09-949-016-5233  
; Sequence 5233, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5233  
; LENGTH: 5925  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5233

Query Match	88.8%	Score	470.8	DB	3	Length	5925
Best Local Similarity	93.0%	Pred. No.	7.3e-140	Mismatches	37	Indels	0
Matches	493	Conservative	0	0	Gaps	0	0

QY 1 GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCAGGAGAAATCGGGGAGGAGAAC 60  
DB 498 GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCAGGAGAAATCGGGGAGGAGAAC 557  
QY 61 TGTGAACACGAGAGCGTGTGATCAGTCTTTGCGCAATCGCGGATCCTGGGAGAGAA 120  
DB 558 TGTGAACACGAGAGCGTGTGATCAGTCTTTGCGCAATCGCGGATCCTGGGAGAGAA 617  
QY 121 GAGGAGGAGAACTTACTCTCGGCAAAATGAGGGGATCTGGAAGATCAGATCATCATGTC 180  
DB 618 GAGGAGGAGAACTTACTCTCGGCAAAATGAGGGGATCTGGAAGATCAGATCATCATGTC 677  
QY 181 CAAACCCCTGCTCGAGGCTTTTGGCAACGCGCAAGACCGTGAGGAGAACGACAACTCCTCTCG 240  
DB 678 CAAACCCCTGCTCGAGGCTTTTGGCAACGCGCAAGACCGTGAGGAGAACGACAACTCCTCTCG 737  
QY 241 CTTTGTGTAATTCATCAGGATCCACTTCGGTACCACTGGGAACTGGCTTCTGCTGACAT 300

```
Db      738 CTTTGGTAAATTCCATCAGGATCCATTCGGTACCAACAGGGAACCTGCTTCTGCTGATAT 797
Qy      301 CGAAACATATCTTCTAGAGAAGTCTAGAGTCACTTTCCAGCTAAAGCAGAAAGAGCTTA 360
Db      798 TGAACACATATCTTCTGAGAAGTCTAGAGTACTTTCCAGCTAAAGGCTGAAAGAGCTA 857
Qy      361 CCACATTTTATCAGATCATGTCTTCAACAAGAGCCAGAGCTCATTTGAAATGCTCCCTGAT 420
Db      858 TCATATTTTATCAGATCATGTCTTCAACAAGAGCCAGAGTCTTAATTTGAAATGCTCCCTGAT 917
Qy      421 CACCACCAACCCATATGACTACGCTTCCTGTCAGTCAAGGGGAGATCACTGTCCCCAGCAT 480
Db      918 CACCACCAACCCATACGATATGCTTCCTGTCAGTCAAGGGGAGATCACTGTCCCCAGCAT 977
Qy      481 TGATGACCAAGAGAGCTGATGGCCACAGATAGTGCATTTGAAATCCTGG 530
Db      978 TGATGACCAAGAGAGTGTGATGGCTTCCGTACAGATAGTGCATTTGAAATCCTGG 1027

RESULT 2
US-09-949-016-1054
; Sequence 1054, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1054
; LENGTH: 6016
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1054

Query Match      80.1%; Score 424.4; DB 3; Length 6016;
Best Local Similarity 87.5%; Pred. No. 5.3e-125;
Matches 464; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy      1 GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGGAAGAC 60
Db      609 GCTAACTGATCGTGAACCAACAGTCAATCTTGATTAATCGGAGAAATCTGGTGCAGGGAAGAC 668
Qy      61 TGTGAACACGAAGCGTGTCTATCCAGTACTTTTGCACAATCGCGTCACTCGGGGAGAGAA 120
Db      669 TGTGAACACGAAGCGTGTCTATCCAGTACTTTTGCACAATTCAGTACTTCTGAGAGAA 728
Qy      121 GAAGGAGGAACCTTACTCTCGGCAAAATGCAAGGAGCTCTGGAAGATCAGATCATCATGTC 180
Db      729 AAAAGAGGAACCTTCTGCGCAAAATGCAAGGAGCCCTTTGAAGATCAAAATCATCATGTC 788
Qy      181 CAACCCCTCTCTCGAGGCTTTGGCAACGCGCAAGACCGTGCAGGAACGACAACTCCTCTCG 240
Db      789 TAACCCCTTACTGGAAGCCTTTCGCAATGCGCAAGCGTGCAGGAATGACAACTCCTCTCG 848
Qy      241 CTTTGGTAAATTCATCAGGATCCACTTCGCTACCAATCGCGTCACTCGGGGAGAGAA 300
Db      849 CTTTGGTAAATTCATCAGGATCCACTTCGCTACCAATCGCGTCACTCGGGGAGAGAA 908
Qy      301 CGAAACATATCTTCTAGAGAAGTCTAGAGTCACTTTCCAGCTAAAGCAGAAAGAGCTTA 360
Db      909 TGAACACATATCTGCTAGAGAAGTCCCGAGTACTTTTCAGCTAAAGGCTGAAAGAGCTA 968
Qy      361 CCACATTTTATCAGATCATGTCTTCAACAAGAGCCAGAGCTCATTTGAAATGCTCCCTGAT 420
Db      969 CACACCAACCCATATGACTTTCGCTTTCAGTCAAGGGGAGATCACTGTCCCCAGCAT 480
Qy      1029 CACCACCAACCCATATGACTTTCGCTTTCAGTCAAGGGGAGAAATTAATCTGTGCCAGCAT 1088
```

```
Db      969 CCACATTTTATCAAAATCTGTCCAATAGAAACACAGAGCTCATTTGAAATGCTTCTGAT 1028
Qy      421 CACCACCAACCCATATGACTTACGCTTCGTCAGTCAAGGGGAGATCACTGTCCCCAGCAT 480
Db      1029 CACCACCAACCCATATGACTTTCGCTTTCAGTCAAGGGGAGAAATTAATCTGTGCCAGCAT 1088
Qy      481 TGATGACCAAGAGAGCTGATGGCCACAGATAGTGCATTTGAAATCCTGG 530
Db      1089 TGATGACCAAGAGAGCTGATGGCCACAGATAGTGTGTCGACATCCTGG 1138

RESULT 3
US-09-949-016-2264
; Sequence 2264, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2264
; LENGTH: 6016
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2264

Query Match      80.1%; Score 424.4; DB 3; Length 6016;
Best Local Similarity 87.5%; Pred. No. 5.3e-125;
Matches 464; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy      1 GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGGAAGAC 60
Db      609 GCTAACTGATCGTGAACCAACAGTCAATCTTGATTAATCGGAGAAATCTGGTGCAGGGAAGAC 668
Qy      61 TGTGAACACGAAGCGTGTCTATCCAGTACTTTTGCACAATCGCGTCACTCGGGGAGAGAA 120
Db      669 TGTGAACACGAAGCGTGTCTATCCAGTACTTTTGCACAATTCAGTACTTCTGAGAGAA 728
Qy      121 GAAGGAGGAACCTTACTCTCGGCAAAATGCAAGGAGCTCTGGAAGATCAGATCATCATGTC 180
Db      729 AAAAGAGGAACCTTCTGCGCAAAATGCAAGGAGCCCTTTGAAGATCAAAATCATCATGTC 788
Qy      181 CAACCCCTCTCTCGAGGCTTTGGCAACGCGCAAGACCGTGCAGGAACGACAACTCCTCTCG 240
Db      789 TAACCCCTTACTGGAAGCCTTTCGCAATGCGCAAGCGTGCAGGAATGACAACTCCTCTCG 848
Qy      241 CTTTGGTAAATTCATCAGGATCCACTTCGCTACCAATCGCGTCACTCGGGGAGAGAA 300
Db      849 CTTTGGTAAATTCATCAGGATCCACTTCGCTACCAATCGCGTCACTCGGGGAGAGAA 908
Qy      301 CGAAACATATCTTCTAGAGAAGTCTAGAGTCACTTTCCAGCTAAAGCAGAAAGAGCTTA 360
Db      909 TGAACACATATCTGCTAGAGAAGTCCCGAGTACTTTTCAGCTAAAGGCTGAAAGAGCTA 968
Qy      361 CCACATTTTATCAGATCATGTCTTCAACAAGAGCCAGAGCTCATTTGAAATGCTCCCTGAT 420
Db      969 CACACCAACCCATATGACTTTCGCTTTCAGTCAAGGGGAGATCACTGTCCCCAGCAT 480
Qy      1029 CACCACCAACCCATATGACTTTCGCTTTCAGTCAAGGGGAGAAATTAATCTGTGCCAGCAT 1088
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:17:34 ; Search time 641.579 Seconds  
(without alignments)  
6831.225 Million cell updates/sec

Title: US-10-785-981-2  
Perfect score: 530  
Sequence: 1 gctgactgacggagaatc.....tagtgcattgaatactctgg 530

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	530	100.0	530	9	US-10-785-981-2
2	530	100.0	530	9	US-10-789-723-2
3	530	100.0	530	9	US-10-786-052-2
4	470.8	88.8	5925	8	US-10-335-053-26
5	432.4	81.6	5874	9	US-10-450-763-21224
6	421.6	79.5	6010	8	US-10-723-860-57
7	421.6	79.5	6010	8	US-10-723-860-105
8	421.6	79.5	6010	8	US-10-723-860-4827
9	421.6	79.5	6010	8	US-10-723-860-4858
10	416.4	78.6	5956	9	US-10-783-271-100
11	362.4	68.4	6523	6	US-10-062-674-1980
12	320.4	60.5	5886	8	US-10-798-037-3
13	320.4	60.5	12801	8	US-10-798-037-5
14	313.2	59.1	5918	5	US-10-198-846-13875
15	308	58.1	4775	7	US-10-336-472-17
16	308	58.1	5780	7	US-10-336-472-19
17	308	58.1	6008	7	US-10-336-472-21
18	308	58.1	6008	7	US-10-717-597-197
19	301.6	56.9	5925	7	US-10-191-803-104
20	248.8	46.9	6149	7	US-10-188-186-95
21	244.2	46.1	716	6	US-10-257-826A-65
22	227.2	42.9	6237	9	US-10-450-763-12161
23	222.4	42.0	6968	10	US-11-097-143-40166

24	162.8	30.7	6915	7	US-10-188-186-47	Sequence 47, Appl
25	143.2	27.0	5896	6	US-10-085-198-131	Sequence 131, App
26	141	26.6	5883	9	US-10-956-157-859	Sequence 859, App
27	141	26.6	6354	5	US-10-084-817-158	Sequence 158, App
28	141	26.6	7274	9	US-10-696-909A-48	Sequence 48, Appl
29	141	26.6	7396	6	US-10-028-248A-35	Sequence 35, Appl
30	141	26.6	7396	7	US-10-107-782-35	Sequence 35, Appl
31	139.2	26.3	2486	8	US-10-357-930-22131	Sequence 22131, A
32	139.2	26.3	2486	8	US-10-357-930-27993	Sequence 27993, A
33	139.2	26.3	2535	3	US-09-927-597-13	Sequence 13, Appl
34	139.2	26.3	2535	9	US-10-486-057-13	Sequence 13, Appl
35	139.2	26.3	6655	7	US-10-296-115-526	Sequence 526, App
36	139.2	26.3	6861	5	US-10-171-311-161	Sequence 161, App
37	139.2	26.3	6861	6	US-10-341-434-102	Sequence 102, App
38	139.2	26.3	6861	9	US-10-923-035-18	Sequence 18, Appl
39	139.2	26.3	6900	5	US-10-171-311-163	Sequence 163, App
40	139.2	26.3	6900	7	US-10-764-425-13	Sequence 13, Appl
41	139.2	26.3	11065	5	US-10-116-802-14	Sequence 14, Appl
42	136.6	25.8	2097	3	US-09-927-597-5	Sequence 5, Appl
43	136.6	25.8	2097	9	US-10-486-057-5	Sequence 5, Appl
44	136.6	25.8	2316	3	US-09-927-597-7	Sequence 7, Appl
45	136.6	25.8	2316	9	US-10-486-057-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-10-785-981-2  
; Sequence 2, Application US/10785981  
; Publication No. US20050112597A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: SCREENING EXPRESSION PROFILE OF GROWTH SPECIFIC GENES IN SWINE  
; FILE REFERENCE: 3884-0120P  
; CURRENT FILING DATE: 2004-06-26  
; PRIOR APPLICATION NUMBER: KR 2003-83653  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-785-981-2

Query Match	100.0%	Score 530;	DB 9;	Length 530;
Best Local Similarity	100.0%	Pred. No. 7.6e-170;	Mismatches 0;	Indels 0;
Matches 530;	Conservative 0;			
Qy	1	GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACGGAGAAATCGGGGAGAAAGAC	60	
Db	1	GCTGACTGATCGGAGAAATCAGTCTATCTTAATCACGGAGAAATCGGGGAGAAAGAC	60	
Qy	61	TGTGAACGAGACGGTGTATCCAGTACTTTGCCAATCCCGTCACTGGGGAGAGAA	120	
Db	61	TGTGAACGAGACGGTGTATCCAGTACTTTGCCAATCCCGTCACTGGGGAGAGAA	120	
Qy	121	GAAGGAGGAACTTACTCTCGGCAAAATGAGGGGACTCTGGAAGATCAGATCATCAGTGC	180	
Db	121	GAAGGAGGAACTTACTCTCGGCAAAATGAGGGGACTCTGGAAGATCAGATCATCAGTGC	180	
Qy	181	CAACCCCTGCTCGAGGCTTTTGGCAACGCGCAAGACGCTGAGGAAAGCAACTCTCTCG	240	
Db	181	CAACCCCTGCTCGAGGCTTTTGGCAACGCGCAAGACGCTGAGGAAAGCAACTCTCTCG	240	
Qy	241	CTTTGGTAAATTCATCAGATCCATTCGGTACCCTGGGAAGCTGCTTCTGCTGACAT	300	
Db	241	CTTTGGTAAATTCATCAGATCCATTCGGTACCCTGGGAAGCTGCTTCTGCTGACAT	300	

Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCAATTTTATCAGATCATGTCTAAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCAATTTTATCAGATCATGTCTAAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
RESULT 2  
US-10-789-723-2  
; Sequence 2, Application US/10789723  
; Publication No. US20050112602A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: cDNA chip for screening specific genes and analyzing their  
; TITLE OF INVENTION: function in swine  
; FILE REFERENCE: YLOP040109US  
; CURRENT APPLICATION NUMBER: US/10/789,723  
; CURRENT FILING DATE: 2004-02-27  
; PRIOR APPLICATION NUMBER: KR 2003-83651  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 530  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-789-723-2  
Query Match 100.0%; Score 530; DB 9; Length 530;  
Best Local Similarity 100.0%; Pred. No. 7.6e-170;  
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTGACTGATCGGAGATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGAAAGAC 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 1 GCTGACTGATCGGAGATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGAAAGAC 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 61 TGTGAACACGAAGCGTGTCTATCTTCAGTACTTTGGCACAATCCCGTCACTGGGAGAAAGAA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 61 TGTGAACACGAAGCGTGTCTATCTTCAGTACTTTGGCACAATCCCGTCACTGGGAGAAAGAA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 121 GAAGGAGAACTTCTCTCGCAAAATGACGGGAGTCTTGGAAGATCAGATCATCAGTGC 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 121 GAAGGAGAACTTCTCTCGCAAAATGACGGGAGTCTTGGAAGATCAGATCATCAGTGC 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 181 CAACCCCTGCTCGAGGCTTTGGCAACGCGCAAGACCGTGAAGACGACAACTCTCTCG 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 181 CAACCCCTGCTCGAGGCTTTGGCAACGCGCAAGACCGTGAAGACGACAACTCTCTCG 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTTGGGAAGCTGGCTTCTGCTGACAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTTGGGAAGCTGGCTTCTGCTGACAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCAATTTTATCAGATCATGTCTAAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCAATTTTATCAGATCATGTCTAAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Qy 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
RESULT 3  
US-10-786-052-2  
; Sequence 2, Application US/10786052  
; Publication No. US20050113568A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: NOVEL GROWTH RELATED GENES FROM SWINE  
; FILE REFERENCE: 3884-0119P  
; CURRENT APPLICATION NUMBER: US/10/786,052  
; CURRENT FILING DATE: 2004-02-26  
; PRIOR APPLICATION NUMBER: KR 2003-83652  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 530  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-786-052-2  
Query Match 100.0%; Score 530; DB 9; Length 530;  
Best Local Similarity 100.0%; Pred. No. 7.6e-170;  
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GCTGACTGATCGGAGATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGAAAGAC 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 1 GCTGACTGATCGGAGATCAGTCTATCTTAATCACCGGAGAAATCCGGGGCAGAAAGAC 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 61 TGTGAACACGAAGCGTGTCTATCTTCAGTACTTTGGCACAATCCCGTCACTGGGAGAAAGAA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 61 TGTGAACACGAAGCGTGTCTATCTTCAGTACTTTGGCACAATCCCGTCACTGGGAGAAAGAA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 121 GAAGGAGAACTTCTCTCGCAAAATGACGGGAGTCTTGGAAGATCAGATCATCAGTGC 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 121 GAAGGAGAACTTCTCTCGCAAAATGACGGGAGTCTTGGAAGATCAGATCATCAGTGC 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 181 CAACCCCTGCTCGAGGCTTTGGCAACGCGCAAGACCGTGAAGACGACAACTCTCTCG 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 181 CAACCCCTGCTCGAGGCTTTGGCAACGCGCAAGACCGTGAAGACGACAACTCTCTCG 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTTGGGAAGCTGGCTTCTGCTGACAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 241 CTTTGGTAAATTCATCAGGATCCACTTCGGTACCACTTGGGAAGCTGGCTTCTGCTGACAT 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 301 CGAAACATATCTTCTAGAGAGTCTAGAGTCACTTTCCAGCTAAAGGAGAGAAAGCTA 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCACATTTTTATCAGATCATGTCTTAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 361 CCACATTTTTATCAGATCATGTCTTAACAAGACCGAGAGTCAATTGAAATGCTCTGAT 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 421 CACCACCAACCATATGACTACGCTTCGTTCAGTCAAGGGGAGATCACTGTCCTCCAGCAT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Qy 481 TGATGACCAAGAGAGTGTGGCCACAGATAGTGCATTGAAATCTCTGG 530  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
RESULT 4  
US-10-335-053-26  
; Sequence 26, Application US/10335053  
; Publication No. US20040241653A1

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:27:30 ; Search time 404.955 Seconds  
(without alignments)  
2778.314 Million cell updates/sec

Title: US-10-785-981-2

Perfect score: 530

Sequence: 1 gctgactgacgggagatc.....tagtgccattgaattcctg 530

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_NA\_New.\*  
1: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/prodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
11: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
12: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	470.8	88.8	5925	12	US-11-000-688-1358
2	416.4	78.6	6062	12	US-11-150-888-13
3	352.8	66.6	6035	12	US-11-136-527-3751
4	328.4	62.0	6021	12	US-11-136-527-274
5	303.6	57.3	5941	12	US-11-136-527-275
6	146	27.5	6060	12	US-11-136-527-2285
7	145.6	27.5	6377	12	US-11-069-834-57
8	145.2	27.4	7355	12	US-11-069-834-47
9	144.8	27.3	6442	12	US-11-069-834-55
10	144.2	27.2	6786	12	US-11-069-834-59
11	141	26.6	7474	12	US-11-069-834-49
12	139.2	26.3	6900	12	US-11-000-688-1142
13	138.8	26.2	4239	8	US-10-995-561-148
14	138.8	26.2	6025	8	US-10-995-561-143
15	138.8	26.2	6148	8	US-10-995-561-147
16	138.8	26.2	6187	8	US-10-995-561-144
17	138.8	26.2	6886	8	US-10-995-561-149
18	138.8	26.2	6925	8	US-10-995-561-145
19	138	26.0	6185	8	US-10-995-561-146
20	135	25.5	2097	12	US-11-136-527-3099

21	130.2	24.6	7666	12	US-11-069-834-51	Sequence 51, Appl
22	126.6	23.9	7853	12	US-11-136-527-2934	Sequence 2934, Ap
23	123.8	23.4	255	8	US-10-909-125-1961	Sequence 1961, Ap
24	122.2	23.1	7619	12	US-11-069-834-53	Sequence 53, Appl
25	109.6	20.7	2529	12	US-11-136-527-3182	Sequence 3182, Ap
26	108.6	20.5	4852	12	US-11-136-527-3350	Sequence 3350, Ap
27	86.6	16.3	5347	9	US-11-072-512-1310	Sequence 1310, Ap
28	85	16.0	7336	12	US-11-136-527-2202	Sequence 2202, Ap
29	84.4	15.9	1568	8	US-10-750-185-37934	Sequence 37934, A
30	84.4	15.9	1568	8	US-10-750-623-37934	Sequence 37934, A
31	83.8	15.8	4456	12	US-11-136-527-318	Sequence 318, App
32	79	14.9	5649	12	US-11-136-527-2615	Sequence 2615, Ap
33	78.8	14.9	636	6	US-09-925-065A-863384	Sequence 863384,
34	78.6	14.8	2727	9	US-11-072-512-1480	Sequence 1480, Ap
35	78	14.7	201	8	US-10-995-561-4989	Sequence 4989, Ap
36	78	14.7	201	8	US-10-995-561-5016	Sequence 5016, Ap
37	78	14.7	201	8	US-10-995-561-5043	Sequence 5043, Ap
38	78	14.7	201	8	US-10-995-561-5073	Sequence 5073, Ap
39	78	14.7	201	8	US-10-995-561-5100	Sequence 5100, Ap
40	78	14.7	201	8	US-10-995-561-5125	Sequence 5125, Ap
41	78	14.7	201	8	US-10-995-561-5144	Sequence 5144, Ap
42	76.6	14.5	1141	8	US-10-750-185-49953	Sequence 49953, A
43	76.6	14.5	1141	8	US-10-750-623-49953	Sequence 49953, A
44	68.8	13.0	538	6	US-09-925-065A-951135	Sequence 951135,
45	68.8	13.0	548	6	US-09-925-065A-944582	Sequence 944582,

#### ALIGNMENTS

#### RESULT 1

US-11-000-688-1358  
; Sequence 1358, Application US/11000688  
; Publication No. US20050287544A1  
; GENERAL INFORMATION:  
; APPLICANT: BERTUCCI, Francois  
; APPLICANT: HOULGATTE, Remi  
; APPLICANT: BIRNBAUM, Daniel  
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS  
; FILE REFERENCE: 1423-R-03  
; CURRENT APPLICATION NUMBER: US/11/000,688  
; PRIOR FILING DATE: 2004-12-01  
; PRIOR APPLICATION NUMBER: US 60/525,987  
; NUMBER OF SEQ ID NOS: 1596  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 1358  
; LENGTH: 5925  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial sequences:primer  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(5925)  
; OTHER INFORMATION: myosin, heavy polypeptide 1, skeletal  
; OTHER INFORMATION: muscle, adult (MvH1) gene.  
US-11-000-688-1358

Query Match	88.8%	Score	470.8	DB	12	Length	5925
Best Local Similarity	93.0%	Pred. No.	7.2e-127				
Matches	493	Conservative	0	Mismatches	37	Indels	0
QY	1	GCTGACTGATCGGAGAAATCAGTCTATCTTATACCGGAGATCCGGGCGGAGGAGAC	60				
Db	498	GCTGACTGATCGGAGAAATCAGTCTATCTTATACCGGAGATCCGGGCGGAGGAGAC	557				
QY	61	TGTGAACAGAAAGCGTGTCTATCTTATACCGGAGATCCGGGCGGAGGAGAA	120				
Db	558	TGTGAACAGAAAGCGTGTCTATCTTATACCGGAGATCCGGGCGGAGGAGAA	617				
QY	121	GAAGGAGGAACCTTCTCTGGGAAAATCAGGGGACTCTGGGAGATCATCATGTC	180				



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:36:14 ; Search time 3563.08 Seconds  
(without alignments)  
9396.598 Million cell updates/sec

Title: US-10-785-981-3  
Perfect score: 589  
Sequence: 1 gttgtcccttaaatatgat.....tagtgcattgaaatcctgg 589

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.4	89.7	1427	6 BD079915	BD079915 Cancer-as
2	528.4	89.7	1677	6 CQ721905	CQ721905 Sequence
3	528.4	89.7	1932	6 CQ848036	CQ848036 Sequence
4	528.4	89.7	2357	6 AX077789	AX077789 Sequence
5	528.4	89.7	2448	8 AK000759	AK000759 Homo sapi
6	528.4	89.7	2703	6 AR338661	AR338661 Sequence
7	528.4	89.7	104089	8 AC012488	AC012488 Homo sapi
8	526.8	89.4	1932	8 AV116969	AV116969 Homo sapi
9	526.8	89.4	2392	8 BC001243	BC001243 Homo sapi
10	498	84.6	2445	8 AB169610	AB169610 Macaca fa
11	498	84.6	2474	8 AB169783	AB169783 Macaca fa
12	466	79.1	184340	14 AC155201	AC155201 Callithri
13	353.8	60.1	776	6 BD101385	BD101385 Novel gen
14	353.8	60.1	776	6 BD021447	BD021447 Novel gen
15	317.8	54.0	517	10 G57776	G57776 SHGC-103708
16	275	46.7	2266	9 BC013617	BC013617 Mus muscu
17	275	46.7	199753	9 AC153650	AC153650 Mus muscu
18	275	46.7	256751	9 AC122925	AC122925 Mus muscu

19	261.2	44.3	308652	14	AC121478	AC121478 Rattus no
20	260	44.1	227221	14	AC096437	AC096437 Rattus no
c 21	234.8	39.9	232244	14	AC114524	AC114524 Rattus no
22	234.8	39.9	238999	14	AC128410	AC128410 Rattus no
23	178	30.2	2316	5	AJ720404	AJ720404 Gallus ga
24	137	23.3	336	6	AR391327	AR391327 Sequence
25	137	23.3	336	6	AR393032	AR393032 Sequence
26	137	23.3	336	6	AR489762	AR489762 Sequence
27	137	23.3	336	6	AR494003	AR494003 Sequence
28	137	23.3	336	6	AX093332	AX093332 Sequence
29	103	17.5	209897	14	AC134122	AC134122 Rattus no
c 30	103	17.5	239398	14	AC105708	AC105708 Rattus no
c 31	94.8	16.1	105208	9	AL671897	AL671897 Mouse DNA
32	82.2	14.0	870	6	BD097742	BD097742 Novel gen
33	82.2	14.0	870	6	BD017804	BD017804 Novel gen
34	80.2	13.6	2167	5	BC051619	BC051619 Dario rer
c 35	78.6	13.3	121251	5	AL591593	AL591593 Zebrafish
36	63.8	10.8	482	10	BV103802	BV103802 MARC 2187
c 37	55.8	9.5	197797	14	AC151536	AC151536 Callithri
38	50	8.5	5866	4	AB025262	AB025262 Sus scrof
c 39	48.8	8.3	2000	6	AX655393	AX655393 Sus scrof
40	48.2	8.2	886	6	AR526567	AR526567 Sequence
41	48.2	8.2	1514	6	AR510239	AR510239 Sequence
42	48.2	8.2	1705	6	CQ597990	CQ597990 Sequence
43	48.2	8.2	1705	6	CQ847838	CQ847838 Sequence
44	48.2	8.2	1705	6	CQ847838	CQ847838 Sequence
45	48.2	8.2	1799	2	AY069456	AY069456 Drosophila

#### ALIGNMENTS

RESULT 1  
BD079915

LOCUS  
BD079915  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Cancer-associated nucleic acids and polypeptides.  
BD079915.1 GI:22625518  
JP 2001516009-A/581.  
Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Mammalia; Eutheria;  
Hominidae; Homo.  
1 (bases 1 to 1427)  
Old, L. J., Scanlan, M. J., Stockert, E., Gure, A., Chen, Y. T., Gout, I.,  
Oghare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahin, U.  
Cancer-associated nucleic acids and polypeptides  
Patent: JP 2001516009-A 581 25-SEP-2001;  
LUDWIG INSTITUTE FOR CANCER RESEARCH  
OS Homo sapiens (human)  
PN JP 2001516009-A/581  
PD 25-SEP-2001  
PF 15-JUL-1998 JP 2000503425  
PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR  
10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR  
11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD  
J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO PI TSENG  
CHEN,  
PI IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFREUNDSCUH, PI  
OZLEM TURECI,  
PI UGUR SAHIN

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1427 bp DNA linear PAT 27-AUG-2002  
Cancer-associated nucleic acids and polypeptides.  
BD079915  
BD079915.1 GI:22625518  
JP 2001516009-A/581.  
Homo sapiens (human)  
Homo sapiens  
Homo sapiens  
Mammalia; Eutheria;  
Hominidae; Homo.  
1 (bases 1 to 1427)  
Old, L. J., Scanlan, M. J., Stockert, E., Gure, A., Chen, Y. T., Gout, I.,  
Oghare, M., Obata, Y., Pfreundschuh, M., Tureci, O. and Sahin, U.  
Cancer-associated nucleic acids and polypeptides  
Patent: JP 2001516009-A 581 25-SEP-2001;  
LUDWIG INSTITUTE FOR CANCER RESEARCH  
OS Homo sapiens (human)  
PN JP 2001516009-A/581  
PD 25-SEP-2001  
PF 15-JUL-1998 JP 2000503425  
PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR  
10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR  
11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD  
J OLD, MATTHEW J SCANLAN, ELISABETH STOCKERT, ALI GURE, YAO PI TSENG  
CHEN,  
PI IVAN GOUT, MICHAEL O'HARE, YUICHI OBATA, MICHAEL PFREUNDSCUH, PI  
OZLEM TURECI,  
PI UGUR SAHIN

FEATURES  
source

Location/Qualifiers  
1. .1427





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:51:20 ; Search time 3489.74 Seconds  
(without alignments)  
7896.735 Million cell updates/sec

Title: US-10-785-981-3  
Perfect score: 589  
Sequence: 1 gttgtcccttaatatgat.....tagtgccattgaaatcctgg 589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.4	89.7	755	2 BG031674	602300827
2	528.4	89.7	1213	4 CR621606	full-leng
3	528.4	89.7	1415	4 CR621025	full-leng
4	528.4	89.7	1512	4 CR594310	full-leng
5	528.4	89.7	1689	4 CR594492	full-leng
6	528.4	89.7	1932	11 DQ053604	Homo sapi
7	527.6	89.6	1012	1 AL560753	AL560753
8	527.4	89.5	1112	3 BM802677	AGENCOURT
9	526.8	89.4	615	2 BE175655	RCS-HT058
10	526.8	89.4	908	2 BG779206	602667157
11	526.8	89.4	1035	1 AL528352	AL528352
12	525.2	89.2	1932	11 DQ053605	DQ053605
13	525	89.1	862	2 BG166586	602340336
14	524	89.0	839	1 AL529078	AL529078
15	517.6	87.9	702	2 BG678319	602626046
16	479.6	81.4	890	1 AL548432	AL548432
17	478.4	81.2	882	2 BF132189	601820139
18	478	81.2	884	2 BF984476	602307829
19	440.4	74.8	485	3 BM766987	K-EST0049
20	428.4	72.7	527	1 AW386704	RCO-PT002
21	419	71.1	604	1 AW581575	HS_3241_A
22	415	70.5	423	9 AQ206754	RCS_3241_A

23	411.4	69.8	507	6	CF134508	CF134508 UI-HF-CB0
24	400.4	68.0	457	2	BE175777	BE175777 RCS-HT058
25	399.6	67.8	767	8	Z78323	Z78323 HSZ78323 Hu
26	385.6	65.5	861	8	DN520090	DN520090 1262794 M
27	384.4	65.3	386	9	AQ207000	AQ207000 HS_3242 A
28	380.4	64.6	587	2	BG403231	BG403231 602418675
29	361.4	61.4	984	7	CO581479	CO581479 ILLUMIGEN
30	360	61.1	757	3	BM990680	BM990680 UI-H-DT0-
31	357.4	60.7	444	2	BE696484	BE696484 PMA-CT039
32	356.8	60.6	382	2	BE773649	BE773649 QV1-FT016
33	353.4	60.0	466	2	BE773939	BE773939 QV1-FT016
34	344	58.4	742	7	CO894304	CO894304 Bowden_22
35	341.8	58.0	912	2	BF213807	BF213807 601847859
36	334.4	56.8	559	1	AW952425	AW952425 EST364495
37	328.4	55.8	812	8	DN872098	DN872098 nad17g11
38	325.4	55.2	759	2	BG164923	BG164923 602343548
39	324.4	55.1	673	2	BE887500	BE887500 601508149
40	317.8	54.0	366	2	BE836524	BE836524 QV0-FN006
41	317.8	54.0	517	9	AQ341682	AQ341682 RPC111-12
42	315.8	53.6	470	1	AJ662831	AJ662831
43	314.4	53.4	716	7	CK835329	CK835329 4059712 B
44	308.2	52.3	384	9	AQ717690	AQ717690 HS_5496 B
45	306.2	52.0	579	2	BE536026	BE536026 601062461

ALIGNMENTS

RESULT 1	BG031674	755 bp	mRNA	linear	EST 24-JAN-2001
LOCUS	602300827F1 NIH_MGC_87 Homo sapiens	CDNA clone	IMAGE:4402182 5',		
DEFINITION	mRNA sequence.				
ACCESSION	BG031674.1	GI:12422197			
VERSION	EST.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 755)				
AUTHORS	NIH-MGC				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: DCTD/DTF CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10110 row: f column: 07 High quality sequence stop: 717.				
FEATURES	Location/Qualifiers				
source	1..755				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:4402182"				
	/tissue_type="mammary adenocarcinoma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_lib="NIH_MGC_87"				
	/note="Organ: Breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."				
ORIGIN					
Query Match	89.7%	Score 528.4;	DB 2;	Length 755;	

Best Local Similarity 98.7%; Pred. No. 1.1e-132; Matches 543; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

QY 1 GTTGTTCCTTTAAATATGATGTTGCCAAGCTGATGGAGACTCATTCGACGTAATATT 60
  |||||
Db 40 GTTGTTCCTTTAAATATGATGTTGCCAAGCTGATGGAGACTCATTCGACGTAATATT 99
  |||||
QY 61 TCCAATGTGCCACTCAAGAGAGATCTTCAAGTCTTTCTTACTGATGTACACATGAAG 120
  |||||
Db 100 TCCAATGTGCCACTCAAGAGAGATCTTCAAGTCTTTCTTACTGATGTACACATGAAG 159
  |||||
QY 121 GAAGTAATTCAGCAGTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTCCTTGTGT 180
  |||||
Db 160 GAAGTAATTCAGCAGTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTCCTTGTGT 219
  |||||
QY 181 TTACCTAGGATGAAACCTGACAGCAAAATGAAGTTTGAAGACGTCGTGTAGGAAAGCA 240
  |||||
Db 220 TTACCTAGGATGAAACCTGACAGCAAAATGAAGTTTGAAGACGTCGTGTAGGAAAGCA 279
  |||||
QY 241 AATGTTGCAATCCTGTTTCTGGGGGCAATGATTCATGCTTATTCGAAACCTTGTCTGAC 300
  |||||
Db 280 AATGTTGCAATCCTGTTTCTGGGGGCAATGATTCATGCTTATTCGAAACCTTGTCTGAC 339
  |||||
QY 301 CGTCATATTCCTTTAGATGAACCAATGATCTTCTTAATGTAGCTTTCATAGCTGAAGAA 360
  |||||
Db 340 CGTCATATTCCTTTAGATGAACCAATGATCTTCTTAATGTAGCTTTCATAGCTGAAGAA 399
  |||||
QY 361 AAGACCATGCCAATCTCTTAAAGATGTTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 420
  |||||
Db 400 AAGACCATGCCAATCTCTTAAAGATGTTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 459
  |||||
QY 421 CTTTCAGAGAAATCTCTAAGATGTTGCTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 480
  |||||
Db 460 CTTTCAGAGAAATCTCTAAGATGTTGCTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 519
  |||||
QY 481 -TCAGTGTACAGATCGAATCAGAGGAGGCGGAGCTAAAGGAACTACAAGCTGTTAGC 539
  |||||
Db 520 GTCAGTGTACAGATCGAATCAGAGGAGGCGGAGCTAAAGGAACTACAAGCTGTTAGC 579
  |||||
QY 540 TGATGACCAA 549
  |||||
Db 580 CCTTCCCGAA 589
  |||||

```

RESULT 2  
 CR621606  
 LOCUS  
 DEFINITION full-length cDNA clone CS0DD004YK21 of Neuroblastoma Cot  
 50-normalized of Homo sapiens (human).  
 CR621606  
 VERSION CR621606.1 GI:50502413  
 KEYWORDS HTC; CNSLT\_CDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominidae; Homo.  
 1 (bases 1 to 1213)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue  
 2 (bases 1 to 1213)  
 Genoscope.  
 Direct Submission  
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)  
 - Web : www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.

FEATURES  
 source  
 Location/Qualifiers  
 1..1213  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DD004YK21"  
 /tissue\_type="Neuroblastoma Cot 50-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 89.7%; Score 528.4; DB 4; Length 1213;  
 Best Local Similarity 98.7%; Pred. No. 1.2e-132;  
 Matches 543; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

QY 1 GTTGTTCCTTTAAATATGATGTTGCCAAGCTGATGGAGACTCATTCGACGTAATATT 60
  |||||
Db 12 GTTGTTCCTTTAAATATGATGTTGCCAAGCTGATGGAGACTCATTCGACGTAATATT 71
  |||||
QY 61 TCCAATGTGCCACTCAAGAGAGATCTTCAAGTCTTCTTACTGATGTACACATGAAG 120
  |||||
Db 72 TCCAATGTGCCACTCAAGAGAGATCTTCAAGTCTTCTTACTGATGTACACATGAAG 131
  |||||
QY 121 GAAGTAATTCAGCAGTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTCCTTGTGT 180
  |||||
Db 132 GAAGTAATTCAGCAGTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTCCTTGTGT 191
  |||||
QY 181 TTACCTAGGATGAAACCTGACAGCAAAATGAAGTTTGAAGACGTCGTGTAGGAAAGCA 240
  |||||
Db 192 TTACCTAGGATGAAACCTGACAGCAAAATGAAGTTTGAAGACGTCGTGTAGGAAAGCA 251
  |||||
QY 241 AATGTTGCAATCCTGTTTCTGGGGGCAATGATTCATGCTTATTCGAAACCTTGTCTGAC 300
  |||||
Db 252 AATGTTGCAATCCTGTTTCTGGGGGCAATGATTCATGCTTATTCGAAACCTTGTCTGAC 311
  |||||
QY 301 CGTCATATTCCTTTAGATGAACCAATGATCTTCTTAATGTAGCTTTCATAGCTGAAGAA 360
  |||||
Db 312 CGTCATATTCCTTTAGATGAACCAATGATCTTCTTAATGTAGCTTTCATAGCTGAAGAA 371
  |||||
QY 361 AAGACCATGCCAATCTCTTAAAGATGTTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 420
  |||||
Db 372 AAGACCATGCCAATCTCTTAAAGATGTTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 431
  |||||
QY 421 CTTTCAGAGAAATCTCTAAGATGTTGCTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 480
  |||||
Db 432 CTTTCAGAGAAATCTCTAAGATGTTGCTGCTGCTGCTGCTGCTGACAGTCCCTAATAACAT 491
  |||||
QY 481 -TCAGTGTACAGATCGAATCAGAGGAGGCGGAGCTAAAGGAACTACAAGCTGTTAGC 539
  |||||
Db 492 GTCAGTGTACAGATCGAATCAGAGGAGGCGGAGCTAAAGGAACTACAAGCTGTTAGC 551
  |||||
QY 540 TGATGACCAA 549
  |||||
Db 552 CCTTCCCGAA 561
  |||||

```

## RESULT 3

CR621025

LOCUS

DEFINITION

full-length cDNA clone CS0DC023YH23 of Neuroblastoma Cot

25-normalized of Homo sapiens (human).

CR621025

VERSION

CR621025.1 GI:50501832

KEYWORDS

HTC; CNSLT\_CDNA.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 1415)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:32:49 ; Search time 499.848 Seconds  
(without alignments)  
7853.391 Million cell updates/sec

Title: US-10-785-981-3

Perfect score: 589

Sequence: 1 gttgttccttaaatatgat.....tagtgccattgaaatcctgg 589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002as.\*

7: Geneseqn2002bs.\*

8: Geneseqn2003as.\*

9: Geneseqn2003bs.\*

10: Geneseqn2003cs.\*

11: Geneseqn2003ds.\*

12: Geneseqn2004as.\*

13: Geneseqn2004bs.\*

14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589	100.0	589	14	Adz89040 Pig growt
2	589	100.0	589	14	Adz79333 Swine gro
3	589	100.0	589	14	Adz77170 Pig growt
4	528.4	89.7	1427	2	Aax40181 Lung canc
5	528.4	89.7	1932	13	Adq89911 Antagonis
6	528.4	89.7	2357	4	Aaf60005 Human syn
7	528.4	89.7	2414	6	Ab189966 Human pol
8	528.4	89.7	2703	4	Aai58275 Human pol
9	528.4	89.7	2703	5	Adq98482 DNA encod
10	528.4	89.7	2703	9	Adb48242 Novel hum
11	528.4	89.7	6499	4	Aak86748 Human imm
12	526.8	89.4	2451	10	Adi40392 Human pur
13	517.4	87.8	2410	4	Aai60061 Human pol
14	474.8	80.6	610	14	ACL57351 Human col
15	391.6	66.5	2734	5	Aas84622 DNA encod
16	353.8	60.1	776	4	Aai97610 Human neu
17	137	23.3	396	4	Aaf94959 Human ova
18	137	23.3	396	6	ABL48909 Ovarian c
19	137	23.3	396	6	ABT03226 Human ova

20	137	23.3	396	11	ADM10819	Adm10819 Human ova	
21	137	23.3	396	12	ADJ11149	Adj11149 Represent	
22	137	23.3	396	12	ADM43410	Adm43410 Human ova	
23	82.2	14.0	870	4	AAI93967	Aai93967 Human neu	
24	59.4	10.1	492	9	ACH21529	Ach21529 Human adu	
25	59	10.0	1208	5	AAS84621	Aas84621 DNA encod	
26	51	8.7	469	14	ADZ89041	Adz89041 Pig growt	
27	51	8.7	469	14	ADZ79334	Adz79334 Swine gro	
28	51	8.7	469	14	ADZ77171	Adz77171 Pig growt	
29	50	8.5	530	14	ADZ89039	Adz89039 Pig growt	
30	50	8.5	530	14	ADZ79332	Adz79332 Swine gro	
31	50	8.5	530	14	ADZ77169	Adz77169 Pig growt	
c	32	48.8	8.3	2000	8	ADA71938	Ada71938 Rice gene
33	48.2	8.2	1705	13	ABL19005	Ab119005 Drosophil	
34	48.2	8.2	1705	13	ADQ89713	Adq89713 Antagonis	
c	35	48.2	8.2	3100	4	ABL18996	Ab118996 Drosophil
c	36	48.2	8.2	3705	4	ABL19004	Ab119004 Drosophil
37	45.4	7.7	540	12	ADK17120	Adk17120 Nancosarch	
c	38	45.4	7.7	9085	12	ADK16049	Adk16049 Continuation (5 of
c	39	44.2	7.5	2000	11	ACL37108	Ac137108 Rice stre
40	43.6	7.4	5925	10	ADD29577	Ad29577 Human tum	
41	43.6	7.4	5925	14	ADM95930	Adm95930 CDNA enco	
42	43.6	7.4	5925	14	ADK15796	Adk15796 DNA encod	
43	43.6	7.4	5925	14	ADK25983	Adk25983 Novel cel	
44	43.6	7.4	6010	12	ADQ22038	Adq22038 Human sof	
45	43.6	7.4	6010	12	ADQ22007	Adq22007 Human sof	

ALIGNMENTS

RESULT 1

ADZ89040

ID ADZ89040 standard; DNA; 589 BP.

XX AC ADZ89040;

XX DT 28-JUL-2005 (first entry)

XX DE Pig growth factor III (GF III) gene, seq id 3.

XX KW Growth; biochip; swine; hog raising; growth factor III; gene; ds.

XX OS Sus scrofa; Kagoshima Berkshire.

XX PN US2005112597-A1.

XX PD 26-MAY-2005.

XX PF 26-FEB-2004; 2004US-00785981.

XX PR 24-NOV-2003; 2003KR-00083653.

XX PA (KIMC//) KIM C.

PA (YEOJ//) YEO J.

PA (LEEJ//) LEE J.

PA (SONG//) SONG Y.

PA (CHOK//) CHO K.

PA (CHUN//) CHUNG K.

PA (KIML//) KIM I.

PA (JINS//) JIN S.

PA (PARK//) PARK S.

PA (JUNG//) JUNG J.

PA (LEEM//) LEE M.

PA (KWON//) KWON E.

PA (CHOE//) CHO E.

PA (CHOH//) CHO H.

PA (SHIN//) SHIN S.

PA (NAMH//) NAM H.

PA (HONG//) HONG Y.

PA (HONG//) HONG S.

PA (KANG//) KANG Y.

PA (HAYY//) HA Y.



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:15:29 ; Search time 144.524 Seconds  
(without alignments)  
7244.353 Million cell updates/sec

Title: US-10-785-981-3

Perfect score: 589

Sequence: 1 gttgttccttaatatgat.....tagtgcattgaaatcctgg 589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PCRUS COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	528.4	89.7	2703	3	US-09-620-312D-152
2	137	23.3	396	3	US-09-640-173-150
3	137	23.3	396	3	US-09-713-550-150
4	137	23.3	396	3	US-09-825-294-150
5	137	23.3	396	3	US-09-970-966-150
6	48.2	8.2	886	3	US-09-270-767-31527
7	48.2	8.2	1514	3	US-09-270-767-15199
8	48	8.1	1141	3	US-09-806-708B-22
9	43.6	7.4	5925	3	US-09-949-016-5233
10	40.4	6.9	6016	3	US-09-949-016-1054
11	40.4	6.9	6016	3	US-09-949-016-2264
12	39.4	6.7	1428	3	US-09-248-796A-1936
13	39.4	6.7	2082	3	US-09-248-796A-2564
14	38.8	6.6	601	3	US-09-949-016-40750
15	38.8	6.6	462589	3	US-09-949-016-12900
16	38.8	6.6	476044	3	US-09-949-016-12412
17	38.6	6.6	723	3	US-09-134-000C-2397
18	38.4	6.5	9048	3	US-08-973-273-4
19	38.2	6.5	1368	3	US-09-248-796A-11323
20	38.2	6.5	1716	3	US-08-656-034-9
21	38	6.5	1698	3	US-09-248-796A-1623
22	37.8	6.4	822	3	US-09-710-279-607
23	37.8	6.4	2950	3	US-09-710-279-3349
24	37.8	6.4	3760	3	US-09-710-279-4029

Sequence 239, App  
Sequence 1000, Ap  
Sequence 1071, Ap  
Sequence 16353, A  
Sequence 1893, Ap  
Sequence 2082, Ap  
Sequence 14502, A  
Sequence 14, Appl  
Sequence 26797, A  
Sequence 2885, Ap  
Sequence 11252, A  
Sequence 15845, A  
Sequence 2570, Ap  
Sequence 2813, Ap  
Sequence 3873, Ap  
Sequence 1, Appli  
Sequence 3017, Ap  
Sequence 145280,  
Sequence 3, Appli  
Sequence 121, App

#### ALIGNMENTS

##### RESULT 1

US-09-620-312D-152  
; Sequence 152, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyang  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pc.fl.genes version 1.0  
; SEQ ID NO 152  
; LENGTH: 2703  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (398)..(2329)  
US-09-620-312D-152

Query Match 89.7%; Score 528.4; DB 3; Length 2703;  
Best Local Similarity 98.7%; Pred. No. 2.3e-146;  
Matches 543; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 GTTGTTCCTTTAAATATGATGTTGCCAAGCTGCATTGCAGACTCATTCAGTAATATT 60



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocorelation Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:17:34 ; Search time 713 Seconds  
(without alignments)  
6831.225 Million cell updates/sec

Title: US-10-785-981-3

Perfect score: 589

Sequence: 1 gttgttccttaaatgat.....tagtgcattgaaatcctgg 589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	589	100.0	589	9	US-10-785-981-3
2	589	100.0	589	9	US-10-789-723-3
3	589	100.0	589	9	US-10-786-052-3
4	528.4	89.7	1932	9	US-10-745-237-341
5	528.4	89.7	2414	6	US-10-264-237-528
6	528.4	89.7	2703	5	US-10-037-270-152
7	528.4	89.7	2703	6	US-10-117-722-152
8	528.4	89.7	2703	9	US-10-122-851-152
9	391.6	66.5	2734	9	US-10-450-763-20426
10	304	51.6	529	4	US-09-925-065A-278866
11	304	51.6	529	4	US-09-925-065A-278867
12	247	41.9	595	4	US-09-925-065A-278868
13	180.2	30.6	510	3	US-09-783-590-2979
14	137	23.3	396	3	US-09-825-294-150
15	137	23.3	396	3	US-09-970-966-150
16	137	23.3	396	6	US-10-212-677-150
17	137	23.3	396	6	US-10-361-811-150
18	137	23.3	396	6	US-10-369-186-150
19	59.4	10.1	492	3	US-09-918-995-8741
20	59	10.0	1208	9	US-10-450-763-20425
21	51	8.7	469	9	US-10-785-981-4
22	51	8.7	469	9	US-10-789-723-4
23	51	8.7	469	9	US-10-786-052-4

24	50	8.5	530	9	US-10-785-981-2	Sequence 2, Appli
25	50	8.5	530	9	US-10-789-723-2	Sequence 2, Appli
26	50	8.5	530	9	US-10-786-052-2	Sequence 2, Appli
27	48.2	8.2	1705	9	US-10-745-237-143	Sequence 143, App
28	48.2	8.2	1705	10	US-11-097-143-25748	Sequence 25748, A
29	48.2	8.2	3100	10	US-11-097-143-25735	Sequence 25735, A
c 29	48.2	8.2	3100	10	US-11-097-143-25747	Sequence 25747, A
c 30	48.2	8.2	3705	10	US-11-097-143-25747	Sequence 25747, A
c 31	45	7.6	637	4	US-09-925-065A-754045	Sequence 754045,
c 32	43.6	7.4	5925	8	US-10-735-053-26	Sequence 26, Appl
c 33	43.6	7.4	6010	8	US-10-723-860-57	Sequence 57, Appl
c 34	43.6	7.4	6010	8	US-10-723-860-105	Sequence 105, App
c 35	43.6	7.4	6010	8	US-10-723-860-4827	Sequence 4827, Ap
c 36	43.6	7.4	6010	8	US-10-723-860-4858	Sequence 4858, Ap
c 37	42	7.1	555	3	US-09-969-034-466	Sequence 466, App
c 38	40.4	6.9	5874	9	US-10-450-763-21224	Sequence 21224, A
c 39	40	6.8	606398	8	US-10-719-993-6782	Sequence 286, App
c 40	40	6.8	18683	6	US-10-311-455-286	Sequence 286, App
c 41	40	6.8	18683	6	US-10-240-452-34	Sequence 34, Appli
c 42	40	6.8	495269	7	US-10-398-221-8	Sequence 8, Appli
c 43	40	6.8	3011208	7	US-10-398-221-2058	Sequence 2058, Ap
c 44	39.8	6.8	64482	7	US-10-322-696-151	Sequence 151, App
c 45	39.6	6.7	8700	6	US-10-311-455-976	Sequence 976, App

ALIGNMENTS

RESULT 1

US-10-785-981-3  
; Sequence 3, Application US/10785981  
; Publication No. US20050112597A1  
; GENERAL INFORMATION:  
; APPLICANT: GYONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: SCREENING EXPRESSION PROFILE OF GROWTH SPECIFIC GENES IN SWINE  
; FILE REFERENCE: 3884-0120P  
; CURRENT FILING DATE: 2004-06-26  
; PRIOR APPLICATION NUMBER: KR 2003-83653  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 3  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-785-981-3

Query Match	100.0%	Score 589;	DB 9;	Length 589;
Best Local Similarity	100.0%	Pred. No. 7.6e-148;		
Matches 589;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GTGTGTCCTTTAAATATGATGTTGCCACAAGCTGCATTGGAGACTCATTCAGTAATATT	60	
Db	1	GTGTGTCCTTTAAATATGATGTTGCCACAAGCTGCATTGGAGACTCATTCAGTAATATT	60	
Qy	61	TCCAATGTGCCACCTCAAGAGAGATCTTCAAGTCTTTCTTACTGATGTACATCAAG	120	
Db	61	TCCAATGTGCCACCTCAAGAGAGATCTTCAAGTCTTTCTTACTGATGTACATCAAG	120	
Qy	121	GAAGTAATTCAGCAGTTCATTGATGCTCCTGAGTGTAGCAGTCAAGAAACGTTCTTGT	180	
Db	121	GAAGTAATTCAGCAGTTCATTGATGCTCCTGAGTGTAGCAGTCAAGAAACGTTCTTGT	180	
Qy	181	TTACCTAGGATGAAACCTCGACAGCAATGAAGTTTGAAGAACTGTGTAGTAAAGCA	240	
Db	181	TTACCTAGGATGAAACCTCGACAGCAATGAAGTTTGAAGAACTGTGTAGTAAAGCA	240	
Qy	241	AATGTTGCATCTCTGTTTCTGGGGCATTGATCCATGCTTATTCGAACCCCTGCTGAC	300	
Db	241	AATGTTGCATCTCTGTTTCTGGGGCATTGATCCATGCTTATTCGAACCCCTGCTGAC	300	





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocorelation Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:27:30 ; Search time 450.034 Seconds  
(without alignments)  
2778.314 Million cell updates/sec

Title: US-10-785-981-3

Perfect score: 589

Sequence: 1 gttgttccttaaatgat.....tagtgccattgaatcctgg 589

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_NA\_New.\*  
1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq2.\*  
11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3.\*  
12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq4.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	304	51.6	529	6	US-09-925-065A-278866
2	304	51.6	529	6	US-09-925-065A-278867
3	247	41.9	595	6	US-09-925-065A-278868
C	4	45	7.6	637	6 US-09-925-065A-754045
5	43.6	7.4	5925	12	US-11-000-688-1358
6	41.2	7.0	2147	8	US-10-750-185-40169
7	41.2	7.0	2147	8	US-10-750-185-40169
C	8	39.4	6.7	493	6 US-09-925-065A-301178
C	9	39.4	6.7	493	6 US-09-925-065A-301179
10	37.8	6.4	822	8	US-10-793-626-607
11	37.8	6.4	1340	12	US-11-136-527-6340
C	12	37.8	6.4	2950	8 US-10-793-626-3349
13	37.8	6.4	3378	12	US-11-136-527-2244
C	14	37.8	6.4	3760	8 US-10-793-626-4029
15	37.4	6.3	405	6	US-09-925-065A-406315
16	37.4	6.3	577	6	US-09-925-065A-474125
17	37.4	6.3	618	6	US-09-925-065A-945465
18	37.4	6.3	643	6	US-09-925-065A-801427
19	37.4	6.3	643	6	US-09-925-065A-801428
20	37.4	6.3	643	6	US-09-925-065A-801429

21	37	6.3	2222	8	US-10-750-185-46406	Sequence 46406, A
22	37	6.3	2222	8	US-10-750-623-46406	Sequence 46406, A
23	36.8	6.2	519	6	US-09-925-065A-207084	Sequence 207084, A
24	36.2	6.1	6035	12	US-11-136-527-3751	Sequence 3751, Ap
25	36.2	6.1	6062	12	US-11-150-888-13	Sequence 13, Appl
C 26	36	6.1	588	6	US-09-925-065A-623204	Sequence 623204, A
C 27	35.8	6.1	587	6	US-09-925-065A-250166	Sequence 250166, A
C 28	35.8	6.1	1660	8	US-10-750-185-50752	Sequence 50752, A
C 29	35.8	6.1	1660	8	US-10-750-623-50752	Sequence 50752, A
C 30	34.8	5.9	577	6	US-09-925-065A-601955	Sequence 601955, A
C 31	34.8	5.9	1922	6	US-09-925-065A-668701	Sequence 668701, A
C 32	34.8	5.9	1922	6	US-09-925-065A-668702	Sequence 668702, A
C 33	34.8	5.9	1922	6	US-09-925-065A-668703	Sequence 668703, A
C 34	34.6	5.9	574	6	US-09-925-065A-339172	Sequence 339172, A
C 35	34.6	5.9	1160	6	US-09-925-065A-951129	Sequence 951129, A
C 36	34.6	5.9	1617	6	US-09-925-065A-673939	Sequence 673939, A
C 37	34.4	5.8	575	6	US-09-925-065A-394695	Sequence 394695, A
C 38	34.4	5.8	2678	6	US-09-925-065A-714439	Sequence 714439, A
C 39	34.4	5.8	2678	6	US-09-925-065A-714440	Sequence 714440, A
C 40	34.2	5.8	1551	8	US-10-750-185-36922	Sequence 36922, A
C 41	34.2	5.8	1551	8	US-10-750-623-36922	Sequence 36922, A
C 42	34.2	5.8	2292	8	US-10-821-234-767	Sequence 767, App
C 43	34.2	5.8	3727	12	US-11-091-883-22	Sequence 22, Appl
C 44	34.2	5.8	3727	12	US-11-091-883-168	Sequence 168, Appl
C 45	34.2	5.8	190276	8	US-10-661-966-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-925-065A-278866  
; Sequence 278866, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 278866  
; LENGTH: 529  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-278866

Query Match 51.8%; Score 304; DB 6; Length 529;  
Best Local Similarity 100.0%; Pred. No. 1.2e-68;  
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GTGTGTCCTTTAAATATGATGTTGCCACAGCTGTCAGTTCAGTCTATTCAGTAAATATT	60
Db	226	GTGTGTCCTTTAAATATGATGTTGCCACAGCTGTCAGTCTATTCAGTAAATATT	285
QY	61	TCCAATGTGCACCTCAAGAGAGATACATTCAGTCTTTCTTACTGATGACATGAAG	120
Db	286	TCCAATGTGCACCTCAAGAGAGATACATTCAGTCTTTCTTACTGATGACATGAAG	345
QY	121	GAAGTAATTCAGAGTTCATTTGATGTCCTGAGTGTAGCAGTCAAGAAACCTGCTTGTGT	180

Db 346 GAAGTAATTCAGCAGTTTCATTGATGCTCCTCAGTGTAGCAGTCAAGAAACGTGCTTGTGT 405  
Qy 181 TTACCTAGGATGAAACCTTGACAGCAAAATGAAGTTTGTGAAACGTTGTGATAGGAAAGCA 240  
Db 406 TTACCTAGGATGAAACCTTGACAGCAAAATGAAGTTTGTGAAACGTTGTGATAGGAAAGCA 465  
Qy 241 AATGTTGCAATCCTGTTTCTGGGGCATTGATCCATGTTTATGCAACCTTGTCTGAC 300  
Db 466 AATGTTGCAATCCTGTTTCTGGGGCATTGATCCATGTTTATGCAACCTTGTCTGAC 525  
Qy 301 CGTC 304  
Db 526 CGTC 529

RESULT 2  
US-09-925-065A-278867  
; Sequence 278867, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 278867  
; LENGTH: 529  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-278867

Query Match 51.6%; Score 304; DB 6; Length 529;  
Best Local Similarity 100.0%; Pred. No. 1.2e-68;  
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTGTTCCTTTAATATGATGTTGCCACAAGCTGCATTGGAGACTCAATTCAGTAATATT 60  
Db 226 GTTGTTCCTTTAATATGATGTTGCCACAAGCTGCATTGGAGACTCAATTCAGTAATATT 285  
Qy 61 TCCAAATGTCACCTCAAGAGAGATACCTCAAGTCTTCTTACTGATGTACACATGAAG 120  
Db 286 TCCAAATGTCACCTCAAGAGAGATACCTCAAGTCTTCTTACTGATGTACACATGAAG 345  
Qy 121 GAAGTAATTCAGCAGTTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTTGTTGT 180  
Db 346 GAAGTAATTCAGCAGTTTCATTGATGTCCTGAGTGTAGCAGTCAAGAAACGTTGTTGT 405  
Qy 181 TTACCTAGGATGAAACCTTGACAGCAAAATGAAGTTTGTGAAACGTTGTGATAGGAAAGCA 240  
Db 406 TTACCTAGGATGAAACCTTGACAGCAAAATGAAGTTTGTGAAACGTTGTGATAGGAAAGCA 465  
Qy 241 AATGTTGCAATCCTGTTTCTGGGGCATTGATCCATGTTTATGCAACCTTGTCTGAC 300  
Db 466 AATGTTGCAATCCTGTTTCTGGGGCATTGATCCATGTTTATGCAACCTTGTCTGAC 525  
Qy 301 CGTC 304  
Db 526 CGTC 529

RESULT 3  
US-09-925-065A-278868  
; Sequence 278868, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 278868  
; LENGTH: 595  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-278868

Query Match 41.9%; Score 247; DB 6; Length 595;  
Best Local Similarity 97.0%; Pred. No. 6.2e-54;  
Matches 261; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 282 TATTGCAACCCCTTGCTGACCGTCATATTCCTTTAGATGAACCAATTGATCTTCTTAATGT 341  
Db 1 TATTGCAACCCCTTGCTGACCGTCATATTCCTTTAGATGAACCAATTGATCTTCTTAATGT 60  
Qy 342 AGCTTTTCATAGCTGAAGAAAGACCATGCCAATCTTTAAACAGAGAAGGGAATAACA 401  
Db 61 AGCTTTTCATAGCTGAAGAAAGACCATGCCAATCTTTAAACAGAGAAGGGAATAACA 120  
Qy 402 GAAATAATAATGTGAATACCTTCAGAAAGATTCCTAAAGATGTTGCTGCTGCTGC 461  
Db 121 GAAATAATAATGTGAATACCTTCAGAAAGATTCCTAAAGATGTTGCTGCTGCTGC 180  
Qy 462 TGACAGTCCCTTAATAAATCAT-TCAAGTGTACAGATCGAATCACAAGGAGGCGGACTAAA 520  
Db 181 TGACAGTCCCTTAATAAATCATGTCAAGTGTACAGATCGAATCACAAGGAGGCGGACTAAA 240  
Qy 521 GGAATCTACAGCTGTTAGCTGATGACCAA 549  
Db 241 GGAATCTACAGCTGTTAGCTGATGACCAA 269

RESULT 4  
US-09-925-065A-754045/c  
; Sequence 754045, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:36:14 ; Search time 2837.15 Seconds  
(without alignments)  
9396.598 Million cell updates/sec

Title: US-10-785-981-4  
Perfect score: 469  
Sequence: 1 cattatgagggtacgcg.....tagtgccattgaaatcctgg 469

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_cm.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	408.2	87.0	1131	8	CR541796
2	408.2	87.0	1134	11	AY893990
3	408.2	87.0	1374	6	CO714735
4	408.2	87.0	1374	6	CS119314
5	408.2	87.0	1381	6	AX834441
6	408.2	87.0	1381	8	AK096902
7	408.2	87.0	1694	8	BC012597
8	406.6	86.7	1134	8	CR536516
9	406.6	86.7	1374	8	HUMACTASK
10	353.2	75.3	1587	5	AB086240
11	352	75.1	1419	9	MUSACSM
12	351.6	75.0	1568	5	AB021652
13	351.6	75.0	1582	5	AB086242
14	350.4	74.7	1457	9	BC014877
15	350	74.6	1611	5	AB021650
16	348.8	74.4	1518	9	BC061974
17	347.2	74.0	1251	9	MMACTASR
18	347.2	74.0	2472	5	AB052654

19	342.4	73.0	1191	9	MUSACSM
20	337.6	72.0	1134	5	AY690421
21	337.6	72.0	1268	5	CRASAA2
22	336	71.6	1325	5	S63494
23	334.4	71.3	1184	5	AY35870
24	334.4	71.3	1188	5	AY35871
25	334.4	71.3	1260	5	CYISAA1
26	334	71.2	1134	5	AF503593
27	330.8	70.5	1134	5	AF503591
28	328	69.9	1134	5	AF303985
29	328	69.9	1330	4	BT021508
30	327.6	69.9	996	5	AY735013
31	327.6	69.9	1140	5	AY35872
32	327.6	69.9	1401	5	AB086241
33	326.4	69.6	1284	5	AF180887
34	326.4	69.6	1302	5	BC045406
35	326.4	69.6	1316	5	BC065435
36	326.4	69.6	1839	4	OCRAGNMA
37	326	69.5	1400	5	AB021651
38	326	69.5	1436	5	AB021649
39	326	69.5	1555	5	AB037866
40	325.2	69.3	1340	9	MMACTASM
41	325.2	69.3	1354	9	BC064800
42	325.2	69.3	1361	6	AX827826
43	325.2	69.3	1361	9	RNACTAV
44	324.8	69.3	1131	2	AY161281
45	324.8	69.3	1134	5	AF503589

#### ALIGNMENTS

#### RESULT 1

#### CR541796

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### COMMENT

CR541796 1131 bp mRNA linear PRI 29-JUN-2004  
Homo sapiens full open reading frame cDNA clone RZPD0834B0631D for  
gene ACTAL, actin, alpha 1, skeletal muscle; complete cds, without  
stopcodon.

CR541796 1 GI:49456548

Full ORF shuttle clone, Gateway(TM), complete cds.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 1131)

Halleck,A., Ebert,L., Mkoondinya,M., Schick,M., Eisenstein,S.,

Neubert,P., Katrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,

Korn,B., Zuo,D., Hu,Y. and LaBaer,J.

Cloning of human full open reading frames in Gateway(TM) system

entry vector (pDONR201)

Unpublished

2 (bases 1 to 1131)

Halleck,A., Ebert,L., Mkoondinya,M., Schick,M., Eisenstein,S.,

Neubert,P., Katrang,K., Schatten,R., Shen,B., Henze,S., Mar,W.,

Korn,B., Zuo,D., Hu,Y. and LaBaer,J.

Direct Submission

Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer

Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,

Germany

RZPD; RZPD0834B0631D, ORFNO 3650

www.rzpd.de/cgi-bin/products/ci.cgi?CloneID=RZPD0834B0631D RZPDLTB;

Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LTB No.

834

www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834

www.rzpd.de/products/orfclones/

Contact: Inge Airlart

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 100

Fax: +49 30 32639 111

www.rzpd.de

This clone is available from RZPD;  
 Contact RZPD (customer.service@rzpd.de) for further information.  
 Clone name at Harvard Institute of Proteomics  
 (www.hip.harvard.edu): FLH130947.01L  
 This CDS clone is part of a collection of human full ORF clones  
 jointly established and verified by the Harvard Institute of  
 Proteomics (HIP) and RZPD.  
 This CDS has been cloned without stopcodon.  
 The CDS has been inserted into pDONR201 via a BP Clonase(TM)  
 reaction. Additional sequence has been added in front of the start  
 codon: att. .AAAAA GCA GGC TCC ACC (ATG).  
 The last codon is followed by the 3' att site: GACCCAGCTTCTT. .att  
 The clone is validated by full sequence check.  
 Compared to the reference sequence NM\_001100 (GI:5016087) we did  
 not find any amino acid exchanges.  
 Clone distribution: http://www.rzpd.de/products/orfclones/.

```

FEATURES
    source
        1..1131
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="RZPD0834B0631D"
            /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
            /lab_host="DH5Alpha"
            /note="vector: pDONR201, Site_1: attP1; Site_2: attP2"
    gene
        1..1131
            /gene="ACTA1"
    CDS
        1..>1131
            /gene="ACTA1"
            /codon_start=1
            /protein_id="CAG46595.1"
            /translation="MCDEBETALVCDNGSLVKAGPAGDAPRAVPFSTVGRPRHOG
            VMVGMGKDSYVGDEAQSKEGILTLKYPTEHGIITWDDMEKIWHHTFYNELRVAPEE
            HPTLLTEAPLNKANREKQIIMFTFNVPMVAIOAVLSVAGRTTTCIGVLDSDGG
            VTHNPVTEGYALPHALRLDLAGRDLDYLMKILTERGYSFVTTAREIVRDIKEKL
            CYVALDFENEMATASSSSLEKSYELPDGQVITIGNERFCRCPETLPQPSFIGMESAGI
            HETYSINMKDCIDIRKDLNANNVMSGTTMYPGIADRMOKREITALAPSTMKIKIAP
            PERKYSVWIGGSILASLSTFQQMWTIKQYEAGPSIVHRKCF"

```

## ORIGIN

```

Query Match      87.0%; Score 408.2; DB 8; Length 1131;
Best Local Similarity 97.9%; Pred. No. 1.6e-58;
Matches 424; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 CATTATTAGGGCTACGCGCTCGCGCAGCATCATGCGCTGACCTGGCGGCGCCGCGA 60
    |||
Db 498 CATTATTAGGGCTACGCGCTCGCGCAGCATCATGCGCTGACCTGGCGGCGCCGCGA 557

Qy 61 TCTCACCAGTACTCGTATGAAGATCTCTCACTGACGCTGGCTACTCTTC-TGACCAACAGC 119
    |||
Db 558 TCTCACCAGTACTCGTATGAAGATCTCTCACTGACGCTGGCTACTCTTCGCTGACCAACAGC 617

Qy 120 TGAGCGCGAGATCTGCGCGACATCAAGGAGACCTGTGCTACGTGGCCCTGGACTTCGA 179
    |||
Db 618 TGAGCGCGAGATCTGCGCGACATCAAGGAGACCTGTGCTACGTGGCCCTGGACTTCGA 677

Qy 180 GAACGAGATGCGGACGCGCGCTCTCTCTCTCTCTCGAAAGAGTACGACGTGCCGAGA 239
    |||
Db 678 GAACGAGATGCGGACGCGCGCTCTCTCTCTCTCGAAAGAGTACGACGTGCCGAGA 737

Qy 240 CGGCGAGTATCATACCATCGGCAACGAGCGCTTCGCTGCCCGGAGAGCGCTCTTCACGCC 299
    |||
Db 738 CGGCGAGTATCATACCATCGGCAACGAGCGCTTCGCTGCCCGGAGAGCGCTCTTCACGCC 797

Qy 300 CTCCTTCATCGGTATGAGTGGCGGCGCATTCACGAGACCACTACACAGCATCATGAA 359
    |||
Db 798 CTCCTTCATCGGTATGAGTGGCGGCGCATTCACGAGACCACTACACAGCATCATGAA 857

Qy 360 GTGTGACATCGACATCAGGAGGACCTGTATGTCACCAACACGTCATGTGCGGGGGCACCAC 419
    |||
Db 858 GTGTGACATCGACATCAGGAGGACCTGTATGTCACCAACACGTCATGTGCGGGGGCACCAC 917

```

```

Qy 420 TGATGACCAAGAG 432
Db 918 GATGTACCCCTGGG 930

```

## RESULT 2

```

AY893990 1134 bp mRNA linear SYN 16-MAR-2005
LOCUS Synthetic construct Homo sapiens clone FLH130947.01L actin alpha 1
DEFINITION (ACTA1) mRNA, partial cds.

```

```

ACCESSION AY893990.1 GI:60832847
VERSION Human ORF Project.
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM 1 (bases 1 to 1134)
REFERENCE Hines, L., Taron, B., Jepson, D., Moreira, D., Raphael, J., Shen, B.,
AUTHORS Halleck, A., Koundinya, M., Hu, Y., Zuo, D. and LaBaer, J.
TITLE Cloning of human full-length CDS FLEXGene in
JOURNAL Gateway(TM) recombinational vector system
REFERENCE 2 (bases 1 to 1134)
AUTHORS Hines, L., Taron, B., Jepson, D., Moreira, D., Raphael, J., Shen, B.,
DIRECT SUBMISSION Halleck, A., Koundinya, M., Hu, Y., Zuo, D. and LaBaer, J.
JOURNAL Submitted (04-JAN-2005) Biological Chemistry and Molecular
Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
Cambridge, MA 02141, USA
COMMENT This CDS clone is a part of a collection of human full-length
expression clones generated by Harvard Institute of Proteomics.
This ORF clone has been cloned without stop-codon (to allow fusion
with C-terminal tag). AttB recombination sites have been added to
either end of the ORF and directionally cloned using the Gateway
cloning system into pDONR 201. Additional sequences in the clone:
'ACC' before the 'ATG' (corresponding to ribosomal binding site and
Kozak consensus sequences). Each clone is clonally isolated and
full-length sequence-verified.

```

```

FEATURES
    Location/Qualifiers
        1..1134
            /organism="synthetic construct"
            /mol_type="mRNA"
            /db_xref="taxon:32630"
            /clone="FLH130947.01L"
            /lab_host="Escherichia coli DH5alpha T1 resistant"
            /note="derived from Homo sapiens first strand cDNA library
            from placenta and brain"
            1..>1134
                /gene="ACTA1"
            1..>1134
                /gene="ACTA1"
                /notes="skeletal muscle"
                /codon_start=1
                /transl_table=11
                /product="actin alpha 1"
                /protein_id="AAK37027.1"
                /db_xref="GI:60832848"
                /translation="MCDEBETALVCDNGSLVKAGPAGDAPRAVPFSTVGRPRHOG
                VMVGMGKDSYVGDEAQSKEGILTLKYPTEHGIITWDDMEKIWHHTFYNELRVAPEE
                HPTLLTEAPLNKANREKQIIMFTFNVPMVAIOAVLSVAGRTTTCIGVLDSDGG
                VTHNPVTEGYALPHALRLDLAGRDLDYLMKILTERGYSFVTTAREIVRDIKEKL
                CYVALDFENEMATASSSSLEKSYELPDGQVITIGNERFCRCPETLPQPSFIGMESAGI
                HETYSINMKDCIDIRKDLNANNVMSGTTMYPGIADRMOKREITALAPSTMKIKIAP
                PERKYSVWIGGSILASLSTFQQMWTIKQYEAGPSIVHRKCF"

```

```

ORIGIN
Query Match      87.0%; Score 408.2; DB 11; Length 1134;
Best Local Similarity 97.9%; Pred. No. 1.6e-58;
Matches 424; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 CATTATTAGGGCTACGCGCTCGCGCAGCATCATGCGCTGACCTGGCGGCGCCGCGA 60
    |||
Db 498 CATTATTAGGGCTACGCGCTCGCGCAGCATCATGCGCTGACCTGGCGGCGCCGCGA 557

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:32:49 ; Search time 398.012 Seconds  
(without alignments)  
7853.391 Million cell updates/sec

Title: US-10-785-981-4

Perfect score: 469

Sequence: 1 cattatgagggtacgcg.....tagtgccattgaaatcctgg 469

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002as.\*

7: Geneseq2002bs.\*

8: Geneseq2003as.\*

9: Geneseq2003bs.\*

10: Geneseq2003cs.\*

11: Geneseq2003ds.\*

12: Geneseq2004as.\*

13: Geneseq2004bs.\*

14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469	100.0	469	14	Adz89041 Pig growt
2	469	100.0	469	14	Adz79334 Swine gro
3	469	100.0	469	14	Adz77171 Pig growt
4	408.2	87.0	1374	14	Adx25976 Novel cel
5	408.2	87.0	1381	11	Adm02880 Human cdn
6	359	76.5	461	8	Abx49264 Bovine ES
7	350.4	74.7	1446	14	Adz60281 Murine Ac
8	347.2	74.0	1134	13	Adv41171 Rat card
9	347.2	74.0	1134	14	Adx26223 Novel cel
10	347.2	74.0	1384	14	Adx25859 Novel cel
11	325.2	69.3	1361	10	Adb53865 Primary r
12	325.2	69.3	1361	11	Adw22181 Rat hepat
13	325.2	69.3	1780	14	Adz60324 Murine Ac
14	312	66.5	1288	12	Adf45416 Human vas
15	312	66.5	1288	12	Adn03844 Antipsori
16	312	66.5	1288	13	Adr24777 Breast ca
17	312	66.5	1288	14	Ady54924 Chronic v
18	312	66.5	1404	4	Aah98660 Human EST
19	308.8	65.8	1275	2	Aat72871 Gamma-smo

20	308.8	65.8	1329	10	ADF76557	Adf76557 Novel hum
21	308.8	65.8	1330	6	ABL62474	ABL62474 Colon ade
22	308.8	65.8	1330	6	ABL69196	ABL69196 Prostate
23	308.8	65.8	1330	6	ABK35532	ABK35532 Gene ACTA
24	308.8	65.8	1330	9	ADB70371	ADB70371 Vascular
25	308.8	65.8	1330	11	ADN95551	ADN95551 Human BEC
26	308.8	65.8	1330	12	ADJ37150	ADJ37150 Human mal
27	308.8	65.8	1330	13	ACN39088	ACN39088 Tumour-as
28	308.8	65.8	1330	14	ADZ26568	ADZ26568 Human smo
29	308.8	65.8	1948	11	ACN90751	ACN90751 Breast ca
30	307.4	65.5	501	6	AAS61617	Aas61617 Lung smal
31	307.2	65.5	561	14	ACL55527	ACL55527 Human col
32	307.2	65.5	1224	8	ACC44347	ACC44347 Gene enco
33	307.2	65.5	1575	6	ABK35123	ABK35123 Human cdn
34	307.2	65.5	1744	12	ADP03041	ADP03041 Human hou
35	307.2	65.5	1744	13	ADS88539	ADS88539 Human hou
36	307.2	65.5	1744	13	ADU60171	ADU60171 Housekeep
37	307.2	65.5	1774	10	ADI62878	ADI62878 Human apo
38	307.2	65.5	1845	13	ADS88823	ADS88823 Nucleotid
39	307.2	65.5	1845	13	ADS88824	ADS88824 Nucleotid
40	307.2	65.5	1845	13	ADS88822	ADS88822 Nucleotid
41	307.2	65.5	1918	6	ABK84502	Abk84502 Human cdn
42	307.2	65.5	1918	14	ADZ26064	Adz26064 Human gen
43	307.2	65.5	1919	6	ABV94253	ABV94253 Breast ca
44	307.2	65.5	1919	7	ADN99914	ADN99914 Human act
45	307.2	65.5	1919	12	ADP03061	ADP03061 Human hou

## ALIGNMENTS

RESULT 1

ADZ89041

ID ADZ89041 standard; DNA; 469 BP.

XX AC ADZ89041;

XX 28-JUL-2005 (first entry)

DE Pig growth factor I (GF IV) gene, seq id 4.

XX Growth; biochip; swine; hog raising; growth factor IV; gene; ds.

XX OS Sus scrofa; Kagoshima Berkshire.

XX PN US2005112597-A1.

XX PD 26-MAY-2005.

XX PF 26-FEB-2004; 2004US-00785981.

XX PR 24-NOV-2003; 2003KR-00083653.

XX PA (KIMC/) KIM C.

PA (YEOJ/) YEO J.

PA (LEEJ/) LEE J.

PA (SONG/) SONG Y.

PA (CHOK/) CHOK K.

PA (CHUN/) CHUNG K.

PA (KIMI/) KIM I.

PA (JINS/) JIN S.

PA (PARK/) PARK S.

PA (JUNG/) JUNG J.

PA (LEEM/) LEE M.

PA (KWON/) KWON E.

PA (CHOE/) CHO E.

PA (CHOH/) CHO H.

PA (SHIN/) SHIN S.

PA (NAMH/) NAM H.

PA (HONG/) HONG Y.

PA (KANG/) KANG Y.

PA (HAYY/) HA Y.



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	408.2	87.0	87.0	483	2	BF826972	BF826972 RC1-HN003
2	408.2	87.0	87.0	524	3	BF769879	BF769879 RC3-IT001
3	408.2	87.0	87.0	583	2	BF316694	BF316694 RC3-IT001
4	408.2	87.0	87.0	872	6	CF552285	CF552285 AGENCOURT
5	408	87.0	87.0	520	8	DR174141	DR174141 HR0001 I2
6	406.6	86.7	86.7	576	5	BX459832	BX459832 DKF2p779J
7	404.6	86.3	86.3	598	3	BF317419	BF317419 RC3-IT001
8	403.4	86.0	86.0	615	7	CR791305	CR791305 DKF3p468B
9	403.4	86.0	86.0	685	7	CR369604	CR369604 170006000
10	403.4	86.0	86.0	1543	4	CR859327	CR859327 Pongo PYG
11	403.2	86.0	86.0	527	2	BF826015	BF826015 MR2-HN003
12	403.2	86.0	86.0	581	3	BP319712	BP319712 RC3-IT001
13	402.4	85.8	85.8	765	7	CR752396	CR752396 DKF3p468B
14	401	85.5	85.5	519	6	CB483530	CB483530 EST_038 S
15	400.2	85.3	85.3	686	6	CD610657	CD610657 56089031H
16	399	85.1	85.1	583	3	BP319352	BP319352 RC3-IT001
17	392.2	83.6	83.6	628	3	BP265037	BP265037 RC3-IT001
18	391.2	83.4	83.4	532	1	AJ710329	AJ710329 AJ710329
19	385.2	82.1	82.1	827	2	BF790986	BF790986 602251034
20	384.2	81.9	81.9	585	1	AA180732	AA180732 ZP43d03.r
21	381.8	81.4	81.4	634	8	DN413152	DN413152 LIB4215-0
22	381	81.2	81.2	526	3	BP265026	BP265026 RC3-IT001

Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 87.0%; Score 408.2; DB 2; Length 483;  
Best Local Similarity 97.9%; Pred. No. 1.2e-89;  
Matches 424; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
Qy 1 CATTATGAGGCTACCGCTGCGCGACGCATCATCGCTGACCTGGCGGGCCGCGA 60  
Db 51 CATTATGAGGCTACCGCTGCGCGACGCATCATCGCTGACCTGGCGGGCCGCGA 110  
Qy 61 TCTCACCAGCTACCTGATGAGATCCTCACTGAGCGTGGCTACTCTTC-TGACCACAGC 119  
Db 111 TCTCACCAGCTACCTGATGAGATCCTCACTGAGCGTGGCTACTCTTCGTGACACAGC 170  
Qy 120 TGACGGGAGATCGTGGCGACATCAAGGAGAGCTGTGCTAGTGGCCCTGGACTTCGA 179  
Db 171 TGACGGGAGATCGTGGCGACATCAAGGAGAGCTGTGCTAGTGGCCCTGGACTTCGA 230  
Qy 180 GAACGAGATGGCGACGCGCGCTCTCTCTCTCTGAAAAGAGCTACGAGCTGCCAGA 239  
Db 231 GAACGAGATGGCGACGCGCGCTCTCTCTCTCTGAAAAGAGCTACGAGCTGCCAGA 290  
Qy 240 CGGCGAGGTATCATACCATCGGCAACGAGCGTTCGCTGCGCCGAGAGCGCTTCTCCAGCC 299  
Db 291 CGGCGAGGTATCATACCATCGGCAACGAGCGTTCGCTGCGCCGAGAGCGCTTCTCCAGCC 350  
Qy 300 CTCCTTCATCGGTATGAGTGGCGGCGATTCAGAGACACCTTACAAAGCATCATGAA 359  
Db 351 CTCCTTCATCGGTATGAGTGGCGGCGATTCAGAGACACCTTACAAAGCATCATGAA 410  
Qy 360 GTGTGATCATGACATCAGGAGGACCTGTATGCAACAGCTCATGTGCGGGGCGACAC 419  
Db 411 GTGTGATCATGACATCAGGAGGACCTGTATGCAACAGCTCATGTGCGGGGCGACAC 470  
Qy 420 TGATGACCAAG 432  
Db 471 GATGATCCCTGGG 483

RESULT 2  
BF769879  
LOCUS RC3-IT0011-161100-022-a03 IT0011 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF769879  
VERSION BF769879.1 GI:12117779  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 524)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-IT0011-161100-022-a03&t3=2000-11-16&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 12  
High quality sequence stop: 524.  
Location/Qualifiers  
1..524  
source

FEATURES  
1..524  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="IT0011"  
/notes="Organ: epid tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 87.0%; Score 408.2; DB 2; Length 524;  
Best Local Similarity 97.9%; Pred. No. 1.2e-89;  
Matches 424; Conservative 0; Mismatches 8; Indels 1; Gaps 1;  
Qy 1 CATTATGAGGCTACCGCTGCGCGACGCATCATCGCTGACCTGGCGGGCCGCGA 60  
Db 37 CATTATGAGGCTACCGCTGCGCGACGCATCATCGCTGACCTGGCGGGCCGCGA 96  
Qy 61 TCTCACCAGCTACCTGATGAGATCCTCACTGAGCGTGGCTACTCTTC-TGACCACAGC 119  
Db 97 TCTCACCAGCTACCTGATGAGATCCTCACTGAGCGTGGCTACTCTTCGTGACACAGC 156  
Qy 120 TGACGGGAGATCGTGGCGACATCAAGGAGAGCTGTGCTAGTGGCCCTGGACTTCGA 179  
Db 157 TGACGGGAGATCGTGGCGACATCAAGGAGAGAGCTGTGCTAGTGGCCCTGGACTTCGA 216  
Qy 180 GAACGAGATGGCGACGCGCGCTCTCTCTCTGAAAAGAGCTACGAGCTGCCAGA 239  
Db 217 GAACGAGATGGCGACGCGCGCTCTCTCTCTGAAAAGAGCTACGAGCTGCCAGA 276  
Qy 240 CGGCGAGGTATCATACCATCGGCAACGAGCGTTCGCTGCGCCGAGAGCGCTTCTCCAGCC 299  
Db 277 CGGCGAGGTATCATACCATCGGCAACGAGCGTTCGCTGCGCCGAGAGCGCTTCTCCAGCC 336  
Qy 300 CTCCTTCATCGGTATGAGTGGCGGCGATTCAGAGACACCTTACAAAGCATCATGAA 359  
Db 337 CTCCTTCATCGGTATGAGTGGCGGCGATTCAGAGACACCTTACAAAGCATCATGAA 396  
Qy 360 GTGTGATCATGACATCAGGAGGACCTGTATGCAACAGCTCATGTGCGGGGCGACAC 419  
Db 397 GTGTGATCATGACATCAGGAGGACCTGTATGCAACAGCTCATGTGCGGGGCGACAC 456  
Qy 420 TGATGACCAAG 432  
Db 457 GATGATCCCTGGG 469

RESULT 3  
BF316694  
LOCUS BP316694 Sugano cDNA library, pericardium Homo sapiens cDNA clone  
DEFINITION BP316694  
ACCESSION BP316694  
VERSION BP316694.1 GI:52245669  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens





Db 815 TGGCAGGTTATACCATGGCAATGAGCGCTTCCGCTGCCGTGAGACCCCTCTTCCAGCC 874  
Qy 300 CTCCTTCATCGGTATGAGTCGGCGGCGCATTCAGGACACCACTTCAACAGCATCATGAA 359  
Db 875 TTCTCTTTATTGGCATGAGTCGCGTGGAAATTCATGAGACCACTTCAAAATTCATCATGAA 934  
Qy 360 GTGTGACATCGACATCAGGAGGACCTGTATGCCAAACAGCTCATGTGCGGGGSCACCAC 419  
Db 935 GTGTGACATTCGATCCGTAGGACTTATATGCCAACATGTCTCTCTCTGCGGGGSCACCAC 994

RESULT 2  
US-09-949-016-4886  
; Sequence 4886, Application US/09949016  
; Patent No. 8812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4886  
; LENGTH: 1288  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-4886

Query Match 66.5%; Score 312; DB 3; Length 1288;  
Best Local Similarity 85.5%; Pred. No. 1.3e-68;  
Matches 359; Conservative 0; Mismatches 60; Indels 1; Gaps 1;  
Qy 1 CATTATGAGGGCTACGCGCTGCGCACGCGCATATGCGCTGGACCTGGCGGCGCCGCGA 60  
Db 575 CATCTATGAAGGCTATGCGCTGCGCCCATGCGCATCATGCGCTGGACTTGGCTGGCGCGTGA 634  
Qy 61 TCTCACCAGTACCTGATGAAGATCTCTCACTGAGGCTGGCTACTCC-TTCTGACCAAGC 119  
Db 635 CTTACGCGACTACCTCATGAAGATCTCTCAGAGAGAGGCTATTCTTTGTGACCAAGC 694  
Qy 120 TGAGCGGAGATGTCGCGGACATCAAGGAGAACTGTGCTACGCTGGCCCTGGACTTGA 179  
Db 695 TGAGAGAGAAATTTGCGAGACATCAAGGAGAACTGTGCTATGTGCGCCCTGGATTTGA 754  
Qy 180 GAACGAGATGCGGAGCGCGCTCTCTCTCTCTCTGAAAAGAGCTACGAGCTGCGCAGA 239  
Db 755 GAATGAGATGCCACAGCAGCTTCTCTCTCTCTCTCTGAGAAGAGCTATGAGCTGCCAGA 814  
Qy 240 CGGGAGGTCATACCAATCGGCAACGAGCGCTTCCGCTGCCGAGAGAGCTCTTCCAGCC 299  
Db 815 TGGCAGGTTATCACCATTGGCAATGAGCGCTTCCGCTGCCCTGAGACCCCTCTTCCAGCC 874  
Qy 300 CTCCTTCATCGGTATGAGTCGGCGGCGCATTCAGGACCACTTCAACAGCATCATGAA 359  
Db 875 TTCTCTTTATTGGCATGAGTCGCGTGGAAATTCATGAGACCACTTCAAAATTCATCATGAA 934  
Qy 360 GTGTGACATCGACATCAGGAGGACCTGTATGCCAAACAGCTCATGTGCGGGGSCACCAC 419  
Db 935 GTGTGACATTCGATCCGTAGGACTTATATGCCAACATGTCTCTCTCTGCGGGGSCACCAC 994

RESULT 3  
US-08-588-113-1  
; Sequence 1, Application US/08588113  
; Patent No. 5710003

; GENERAL INFORMATION:  
; APPLICANT: McHugh, Kirk M.  
; TITLE OF INVENTION: DIAGNOSTIC TEST FOR DETERMINING  
; MALIGNANCY OF SMOOTH MUSCLE TUMORS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 5710003ris  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,113  
; FILING DATE:  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ralph, Rebecca L.  
; REGISTRATION NUMBER: 35,152  
; REFERENCE/DOCKET NUMBER: TJU-1652  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1275 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 55..1186  
US-08-588-113-1

Query Match 65.8%; Score 308.8; DB 2; Length 1275;  
Best Local Similarity 85.0%; Pred. No. 7.9e-68;  
Matches 357; Conservative 0; Mismatches 62; Indels 1; Gaps 1;  
Qy 1 CATTATGAGGGCTACGCGCTGCGCACGCGCATATGCGCTGGACCTGGCGGCGCCGCGA 60  
Db 549 CATCTATGAAGGCTATGCGCTGCGCCCATGCGCATCATGCGCTGGACTTGGCTGGCGCGTGA 608  
Qy 61 TCTCACCAGTACCTGATGAAGATCTCTCACTGAGCGTGGCTACTCC-TTCTGACCAAGC 119  
Db 609 CCTCAGCGACTACTCTCATGAAGATCTCTCAGAGAGAGGCTATTCTTTGTGACCAAGC 668  
Qy 120 TGAGCGCGAGATCGTCGCGGACATCAAGGAGAAAGCTGTGCTACGTGGCCCTGGACTTGA 179  
Db 669 TGAGAGAGAAATTTGCGGAGACATCAAGGAGAGCTGTGCTATGTGCGCCCTGGATTTGA 728  
Qy 180 GAACGAGATGCGGAGCGCGCTCTCTCTCTCTCTGAAAAGAGCTACGAGCTGCGCAGA 239  
Db 729 GAATGAGATGCCACAGCAGCTTCTCTCTCTCTCTGAGAAGAGCTATGAGCTGCGCAGA 788  
Qy 240 CGGGCAGGTCATCACCATCGGCAACGAGCGCTTCCGCTGCCGAGAGAGCTCTTCCAGCC 299  
Db 789 TGGCGAGGTTATCACCATTGGCAATGAGCGCTTCCGCGTCCCTGAGAACCTCTTCCAGCC 848  
Qy 300 CTCCTTCATCGGTATGAGTCGGCGGCGCATTCAGGACCACTTCAACAGCATCATGAA 359  
Db 849 TTCTCTTTATTGGCATGAGTCGCGTGGAAATTCATGAGACCACTTCAAAATTCATCATGAA 908  
Qy 360 GTGTGACATCGACATCAGGAGGACCTGTATGCCAAACAGCTCATGTGCGGGGSCACCAC 419  
Db 909 GTGTGACATTCGATCCGTAGGACTTATATGCCAACATGTCTCTCTCTGCGGGGSCACCAC 968

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:17:34 ; Search time 567.737 Seconds  
(without alignments)  
6831.225 Million cell updates/sec

Title: US-10-785-981-4

Perfect score: 469

Sequence: 1 catttatggggctacgcgc.....tagtgccattgaatctcgg 469

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/prodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	469	100.0	469	9	US-10-785-981-4
2	469	100.0	469	9	US-10-789-723-4
3	469	100.0	469	9	US-10-786-052-4
4	408.2	87.0	1381	6	US-10-108-260A-1565
5	406.6	86.7	1374	7	US-10-240-425-1260
6	359	76.5	461	3	US-09-960-352-14429
7	350.4	74.7	1446	9	US-10-764-420-5
8	347.2	74.0	1134	7	US-10-191-803-116
9	325.2	69.3	1361	6	US-10-388-934-560
10	325.2	69.3	1780	9	US-10-764-420-48
11	312	66.5	1288	6	US-10-341-434-100
12	312	66.5	1288	7	US-10-172-118-638
13	312	66.5	1288	7	US-10-342-887-638
14	312	66.5	1288	9	US-10-923-035-17
15	308.8	65.8	1330	3	US-09-969-708-62
16	308.8	65.8	1330	3	US-09-873-367C-811
17	308.8	65.8	1330	6	US-10-236-031B-63
18	308.8	65.8	1330	8	US-10-370-715B-231
19	308.8	65.8	1330	9	US-10-843-641A-811
20	308.8	65.8	1330	9	US-10-843-641A-7533
21	308.8	65.8	1948	5	US-10-198-846-11901
22	307.4	65.5	501	3	US-09-833-790-158
23	307.2	65.5	1400	9	US-10-956-157-5886

#### ALIGNMENTS

##### RESULT 1

US-10-785-981-4  
; Sequence 4, Application US/10785981  
; Publication No. US20050112597A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: SCREENING EXPRESSION PROFILE OF GROWTH SPECIFIC GENES IN SWINE  
; TITLE OF INVENTION: AND FUNCTIONAL CDNA CHIP PREPARED BY USING THE SAME  
; FILE REFERENCE: 3884-0120P  
; CURRENT APPLICATION NUMBER: US/10/785,981  
; CURRENT FILING DATE: 2004-06-26  
; PRIOR APPLICATION NUMBER: KR 2003-83653  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 4  
; LENGTH: 469  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-785-981-4

Query Match		100.0%;	Score 469;	DB 9;	Length 469;
Best Local Similarity		100.0%;	Pred. No. 2e-128;		
Matches 469;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CATTATGAGGGCTACGCGCTGCGCAGCGCATCATGCGCTGGAGCTGCGGCGCGCGA	60		
Db	1	CATTATGAGGGCTACGCGCTGCGCAGCGCATCATGCGCTGGAGCTGCGGCGCGCGA	60		
QY	61	TCTCACCGACTACTCTGATGAAGATCTCTACTGAGCTGGCTACTCTCTTGACACAGCT	120		
Db	61	TCTCACCGACTACTCTGATGAAGATCTCTACTGAGCTGGCTACTCTCTTGACACAGCT	120		
QY	121	GAGCGGAGATCGTGGCGGACATCAAGGAGAAGCTGTGCTACGTGCGCTTGGACTTCGAG	180		
Db	121	GAGCGGAGATCGTGGCGGACATCAAGGAGAAGCTGTGCTACGTGCGCTTGGACTTCGAG	180		
QY	181	ACGAGATGGGAGCGCGCGCTCTCTCTCTCTGCGGAAAGAGCTACGAGCTGCCAGAC	240		
Db	181	ACGAGATGGGAGCGCGCGCTCTCTCTCTCTGCGGAAAGAGCTACGAGCTGCCAGAC	240		
QY	241	GGCAGGTATCACCATCGGACGAGCTTCGCTGCGCGGAGAGCTCTTCCAGCCC	300		
Db	241	GGCAGGTATCACCATCGGACGAGCTTCGCTGCGCGGAGAGCTCTTCCAGCCC	300		

Sequence 261, App  
Sequence 132, App  
Sequence 244, App  
Sequence 166, App  
Sequence 342, App  
Sequence 856, App  
Sequence 651, App  
Sequence 14, App1  
Sequence 345, App  
Sequence 14002, A  
Sequence 2906, Ap  
Sequence 157, App  
Sequence 393, App  
Sequence 32, App1  
Sequence 156, App  
Sequence 16735, A  
Sequence 16736, A  
Sequence 12, App1  
Sequence 728, App  
Sequence 15488, A  
Sequence 63547, A



1	CATTATTAGAGGGCTACGCGTCCGCCACGCCCATCATGCGCTGACCTGCGCGGCCCGCA	60	Qy
601	CATTATTAGAGGCTACGCGTCCGCCACGCCCATCATGCGCTGACCTGCGCGGCCCGCA	660	Db
61	TCTCCACGACTTACTGATGAAGATCCTCATGAGCGTGGCTATCTCTTC-TGACCA	119	Qy
661	TCTCACCCACTTACTGATGAAGATCCTCATGAGCGTGGCTATCTCTCTGTGACCA	720	Db
120	TGAGCGGAGATCGTGGCCGCAATCAAGGAGAACTGTGTCTACGTGGCCCTG	179	Qy
721	TGAGCGGAGATCGTGGCCGCAATCAAGGAGAACTGTGTCTACGTGGCCCTG	780	Db



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:36:14 ; Search time 3067.03 Seconds  
(without alignments)  
9396.598 Million cell updates/sec

Title: US-10-785-981-5

Perfect score: 507

Sequence: 1 tatatagaaccgaatcacgt.....tctgtcggttcagccgcag 507

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5893141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_cm.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	42.4	8.4	2000	6	AX655393	AX655393 Sequence
C 2	41.4	8.2	4001	2	DRODROSOPH	M23391 Drosophila
C 3	41	8.1	110000	1	CP000035_06	Continuation (7 of
C 4	41	8.1	110000	1	CP000035_07	Continuation (8 of
C 5	40.4	8.0	79951	8	HSJ908M14	AL121832 Human DNA
C 6	39.8	7.9	2852	6	CQ573420	Sequence
C 7	39.8	7.9	3176	2	AY058690	AY058690 Drosophil
C 8	39.8	7.9	5023	6	CQ573419	Sequence
C 9	39.8	7.9	37841	14	AC017708	AC017708 Drosophil
C 10	39.8	7.9	168471	2	AC006170	AC006170 Drosophil
C 11	39.8	7.9	231562	2	AB003767	AB003767 Drosophil
C 12	39.6	7.8	2014	8	HSU82827	U82827 Human trans
C 13	39.6	7.8	2453	8	BC075791	BC075791 Homo sapi
C 14	39.6	7.8	110000	15	AP008214_200	Continuation (201
C 15	39.6	7.8	129355	8	AC004080	AC004080 Homo sapi
C 16	39.6	7.8	176120	5	AP005521	AP005521 Oryza sat
C 17	39.4	7.8	3512	15	AK110440	AK110440 Oryza sat
C 18	39	7.7	7218	6	I66494	I66494 Sequence 14

19	38.8	7.7	339	2	AY333202	AY333202 Anopheles
20	38.8	7.7	177086	5	BX470219	BX470219 Zebrafish
21	38.6	7.6	110000	15	AE016819_06	Continuation (7 of
22	38.6	7.6	110000	15	AE016819_07	Continuation (8 of
23	38.6	7.6	214503	14	AC161180	AC161180 Mus muscu
C 24	38.6	7.6	349895	1	BX248359	BX248359 Corynebac
C 25	38.4	7.6	381	2	AY333201	AY333201 Anopheles
C 26	38.4	7.6	110000	1	CR543861_25	Continuation (26 o
C 27	38.4	7.6	210950	9	AC129317	AC129317 Mus muscu
C 28	38	7.5	96081	15	AP004328	AP004328 Oryza sat
C 29	38	7.5	110000	15	AP008212_190	Continuation (191
C 30	38	7.5	158857	15	AP005470	AP005470 Oryza sat
C 31	38	7.5	197714	14	AC128967	AC128967 Rattus no
C 32	37.6	7.4	221952	14	AC098902	AC098902 Rattus no
C 33	37.6	7.4	256796	14	AC099183	AC099183 Rattus no
C 34	37.4	7.4	4186	1	AF199442	AF199442 Methanoco
C 35	37.4	7.4	34892	15	AC156447	AC156447 Volvox ca
C 36	37.2	7.3	230337	9	AC094571	AC094571 Rattus no
C 37	37.2	7.3	233270	14	AC094486	AC094486 Rattus no
C 38	37.2	7.3	347050	1	AL591981	AL591981 Listeria
C 39	37.2	7.3	349980	6	AX641670	AX641670 Sequence
C 40	36.8	7.3	27561	2	CEY66A7AR	AL590342 Caenorhab
C 41	36.8	7.3	42287	14	AC151614	AC151614 Emiliana
C 42	36.8	7.3	43395	2	CEY66A7A	Continuation (11 o
C 43	36.8	7.3	110000	1	AP006627_10	Continuation (2 of
C 44	36.8	7.3	110000	1	BX571966_01	CR361548 Zebrafish
C 45	36.8	7.3	146409	5	CR361548	

#### ALIGNMENTS

RESULT 1  
AX655393/C  
LOCUS AX655393 Sequence 5263 from Patent WO03000898.  
DEFINITION AX655393  
ACCESSION AX655393  
VERSION AX655393.1 GI:29158207  
KEYWORDS  
SOURCE Oryza sativa  
ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1  
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;  
SYNGENTA PARTICIPATIONS AG (CH)  
FEATURES  
Location/Qualifiers  
source 1..2000  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

ORIGIN  
Query Match 8.4%; Score 42.4; DB 6; Length 2000;  
Best Local Similarity 11.2%; Pred. No. 0.44;  
Matches 40; Conservative 161; Mismatches 157; Indels 0; Gaps 0;

QY	146	CATCCCTTCTGCTGCTGATGTTGGCGTGCAGTTATAGGCTGCCAACCCGCATACA	205
DB	449	YMWSMYKCRSMKYGAKCYGCKMWTYCSYGYMKWTYNGSYKSRCYKMYMYKMW	390
QY	206	CACCTTACAGGCACCTATTAAAGTTATACACAGAGGCTCTTACACCCCTAAGCAGT	265
DB	389	MYMYSYSSMMTWTYYAKYKWKYKRRGTMSTYKSKYKYYCTWVCYMKCMRCYRWR	330
QY	266	GGCAGTGGTAGCGCTGCCCGCTTACCTGCCGAGTGTGGTCTAGCTCCGCTTAAGC	325
DB	329	KWMRKTKTKRCYCWCYRATCYATCWCCCYRKRGRGYSRSMRTAGKWKWRSWRCYSWY	270

```
Qy 326 TTCCCGATAGCCGCGCTTTTACACACCATCGCGGAGTACACACCGCTGGTTCGAGC 385
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 269 KMYKMKWKKSYMSYGHARSSTGWSRAARTYKGYSTSRAXMWRACRMYSACRYSRT 210
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Qy 386 GTAAGCGCTATGATAGACAGCTCGCGGACGCGCGTGTAGCCAGCTTACTACATGTTAGT 445
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 209 SYTCGCSYCGSSKWKYMSKSCMRMTCSSCCYTCYTGAMCWSCCMSMMYMGSCGYT 150
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Qy 446 TTCAGCAACACCTCGCAATACCGGTGTTCCCTACTCCAACTCTGTCGGTTTCAGCC 503
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db 149 RGMKRSKYSMCKCKYCSCCTKYCSYTGYYRYCKWYKYSYKCYCYCYWYMSYMYM 92
    : : : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :

RESULT 2
DRODROSOPH/c DRODROSOPH 4001 bp DNA linear INV 17-AUG-1993
LOCUS Drosophila melanogaster serendipity (sry h-1) gene, complete cds.
DEFINITION M23391
ACCESSION M23391
VERSION M23391.1 GI:340903
KEYWORDS DNA-binding protein; multifingered protein gene.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 4001)
Vincent,A., Kejzlarova-Lepesant,J., Segalat,L., Yanicostas,C. and
Lepesant,J.A.
sry h-1, a new Drosophila melanogaster multifingered protein gene
showing maternal and zygotic expression
Mol. Cell. Biol. 8 (10), 4459-4468 (1988)
3141791
Original source text: Drosophila melanogaster (strain Canton S)
(library: Lauer and Maniatis) DNA.
Location/Qualifiers
1..4001
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Canton S"
/db_xref="taxon:7227"
/map="region 98EF"
/tissue_lib="Lauer and Maniatis"
join(399..1661,1722..2926,3037..3751)
/gene="sry h-1"
join(399..1661,1722..2926,3037..3751)
/gene="sry h-1"
/note="mRNA including exon 1a"
399..1661
/gene="sry h-1"
/note="version a of exon 1"
number=1
join(410..1661,1722..2926,3037..3751)
/gene="sry h-1"
/note="mRNA including exon 1b"
410..1661
/gene="sry h-1"
/note="version b of exon 1"
number=1
join(419..1661,1722..2926,3037..3751)
/gene="sry h-1"
/note="mRNA including exon 1c"
419..1661
/gene="sry h-1"
/note="version c of exon 1"
number=3
join(649..1661,1722..2926,3037..3425)
/gene="sry h-1"
/standard_name="serendipity"
/codon_start=1
/product="DNA-binding protein"
/protein_id="AAA28487.1"
/db_xref="GI:387593"
/translation="MEGKGGRKMBEAPSKLPPIYGGDAGTPTKAAHDEITLSL
```

```
LRINPDSISSIKDESLLIDLACVTTISSASLVNGNSLSSTDFRVLDESAAQNNTLNL
LSQVDRDDLAASTSTLSTLTSDNHSSEFSVTFLRPPNPATNTPPKTSSSGT
SPVKLRSQOLHQHOLQWQSQOLQKPLPAATAVRLKVPKEEPEEKHPPEOVVT
KYVCSSELLPFSFTIFQQAQKSAESVADRLACHPAASETKPLEVDPAPLHKLDGGL
LUETDEVAAGHEAAHRLTLTYRCSQREFELLAGLKKHLKTHRTGREGKDWKCKPD
CGKCLKGSMWHRKIHSDNKKYQCDICQKQFVKQNLTHHARIHSSEKPYECPEQCK
RFQERSHLQHQHYHAQTSYRCEKCGMYKTERCLKVHNLVHLEQRPACTVCDKSF
ISNKLQHSNTHTGMRPFKNCYCPDRFTNFPNWLKHTRRHKVDRHKTGHEHNIIPSY
CSKSTNTKAQAAAAAASAVNPNELSSASSELKAKANLTSTAAPAKAKOARK
KQPOQATLAALGITLPACTALQOHPHPVLAQHQHQLTTLVPLAPPKOTKAKEE
RKQLAPLQKQKQQLQOQOQSSLEPIPAVPOIKKEFPVQGTGPGFDLHLGLSITAE
ELIMEQALACPSLESSPSSPASMELTSDNAISDSAAALHFOIKNELPDLLEDDFLA
PCKPSDLACPSLESSPSSPASMELTSDNAISDSAAALHFOIKNELPDLLEDDFLA
HGKLSSTGNGVQSVTTTSLAQTPSVSMVNVPVLLVRSNQMLPSVDTLTFTNQTGSRFFA
GKSATAATPHLT"
1662..1721
/gene="sry h-1"
/number=1
1722..2926
/gene="sry h-1"
/number=2
2927..3036
/gene="sry h-1"
/number=2
3037..3751
/gene="sry h-1"
/number=3
3728..3733
/gene="sry h-1"

intron
exon
intron
exon
polyA_signal
ORIGIN
Query Match 8.2%; Score 41.4; DB 2; Length 4001;
Best Local Similarity 54.2%; Pred. No. 0.89;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 258 TAAGCAGTGGCAGTGGTAGCCGCTTACCGCTTACCTGCGCAGTGTGTGTGCTAGCTCCG 317
Db 1345 TAAACGATGCGCGCAGTAGCTCAGACTCGCACACCTTACCTTGGTGACCACTTGTCTCCG 1286
Qy 318 TCCTAGCTTCCCGATAGCGCGGCTTTTACACACCATCGCGGACTAGACCGTGTG 377
Db 1285 GCGGATGCTTCTCTCTCGGCGGCTCTCTCTTAAAGACCTTCAGACGCCCGCTCCCG 1226
Qy 378 GTTGACGCGTAAGCGTCTATGTGTAGCAGCTGCGGC 412
Db 1225 CCGCAGCTTGTGCTTCTGCTTGACGAGCTGAGAC 1191
```

RESULT 3  
CP000095.06  
WPCOMMENT

Sequence split into 19 fragments LOCUS CP000095 Accession CP000095

Fragment Name	Begin	End
CP000095_00	1	110000
CP000095_01	100001	210000
CP000095_02	200001	310000
CP000095_03	300001	410000
CP000095_04	400001	510000
CP000095_05	500001	610000
CP000095_06	600001	710000
CP000095_07	700001	810000
CP000095_08	800001	910000
CP000095_09	900001	1010000
CP000095_10	1000001	1110000
CP000095_11	1100001	1210000
CP000095_12	1200001	1310000
CP000095_13	1300001	1410000
CP000095_14	1400001	1510000
CP000095_15	1500001	1610000
CP000095_16	1600001	1710000
CP000095_17	1700001	1810000
CP000095_18	1800001	1842899

Continuation (7 of 19) of CP000095 from base 600001 (CP000095 Prochlorococcus marinus st



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:32:49 ; Search time 430.26 Seconds  
(without alignments)  
7853.391 Million cell updates/sec

Title: US-10-785-981-5

Perfect score: 507

Sequence: 1 tatatagaaccgaatcacgt.....tctgtcggttcagccgcag 507

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	507	100.0	507	14	ADZ89042 Pig growt
2	507	100.0	507	14	ADZ79335 Swine gro
3	507	100.0	507	14	ADZ77172 Pig growt
C 4	42.4	8.4	2000	8	ADA71938 Rice gene
C 5	41.4	8.2	336	2	AAX200010 Antifreez
C 6	41.4	8.2	336	2	AAX88092 Synthetic
C 7	40.8	8.0	891	11	ACL30038 Rice abio
C 8	39.8	7.9	336	2	AAX88093 Synthetic
C 9	39.8	7.9	2852	4	ABL02625 Drosophil
10	39.8	7.9	5023	4	ABL02624 Drosophil
11	39.6	7.8	1167	6	ABK32192 Prostata
12	39.6	7.8	5643	6	ABK931133 Human pro
C 13	37.2	7.3	2712	8	ACA36793 Prokaryot
C 14	37.2	7.3	110000	6	ABAQ3041_18 Continuation (19 o
15	36.8	7.3	362	14	ADV77137 Huntingto
16	36.8	7.3	204803	12	ADQ97348 Mouse can
C 17	36	7.1	728	6	ABQ68881 Listeria
C 18	36	7.1	1296	6	ABQ67797 Listeria
19	36	7.1	1311	6	ABQ67800 Listeria

C 20	36	7.1	2034	6	ABQ70741 Listeria
C 21	36	7.1	110000	4	AAI99683_16 Continuation (17 o
C 22	36	7.1	110000	6	ABQ69245_19 Continuation (20 o
C 23	36	7.1	319630	6	ABQ67194 Listeria
24	35.4	7.0	539	13	ADX10450 Plant ful
25	35.4	7.0	561	13	ADX34527 Plant ful
26	35.4	7.0	642	13	ADX14288 Plant ful
27	35.4	7.0	856	12	ADM47617 Polynucle
C 28	35	6.9	192	13	ADU02160 Novel hum
C 29	35	6.9	863	6	ABK77958 Bacillus
C 30	35	6.9	2511	10	ADJ92192 Human hai
C 31	35	6.9	9732	5	AAS79690 DNA encod
32	34.8	6.9	3990	14	AE91810 DNA encod
C 33	34.8	6.9	110000	4	AAI99682_16 Continuation (17 o
C 34	34.8	6.9	218802	14	ADW98820 Human her
35	34.6	6.8	336	2	AAX200010 Antifreez
36	34.6	6.8	336	2	AAX88092 Synthetic
C 37	34.6	6.8	408	8	ACA38339 Prokaryot
C 38	34.6	6.8	411	8	ACA40379 Prokaryot
C 40	34.6	6.8	1010	5	AAS68262 DNA encod
C 41	34.6	6.8	1167	8	ACA54136 Prokaryot
C 42	34.6	6.8	2466	6	ABS69975 Aspergill
C 43	34.6	6.8	110000	4	AAI99682_07 Continuation (8 of
C 44	34.6	6.8	110000	4	AAI99682_08 Continuation (9 of
C 45	34.6	6.8	110000	4	AAI99683_07 Continuation (8 of
					AAI99683_08 Continuation (9 of

ALIGNMENTS

RESULT 1

ADZ89042

ID ADZ89042 standard; DNA; 507 BP.

XX AC ADZ89042;

XX AC ADZ89042;

28-JUL-2005 (first entry)

Pig growth factor I (GF V) gene, seq id 5.

Growth; biochip; swine; hog raising; growth factor V; gene; ds.

Sus scrofa; Kagoshima Berkshire.

US2005112597-A1.

26-MAY-2005.

26-FEB-2004; 2004US-00785981.

24-NOV-2003; 2003KR-00083653.

(KIMC/) KIM C.

(YEOJ/) YEO J.

(LEEJ/) LEE J.

(SONG/) SONG Y.

(CHOK/) CHO K.

(CHUN/) CHUNG K.

(KIML/) KIM I.

(JINS/) JIN S.

(PARK/) PARK S.

(JUNG/) JUNG J.

(LEEM/) LEE M.

(KWON/) KWON E.

(CHOE/) CHO E.

(CHOH/) CHO H.

(SHIN/) SHIN S.

(NAMH/) NAM H.

(HONG/) HONG Y.

(HONG/) HONG S.

(KANG/) KANG Y.

(HAYY/) HA Y.

```

PA (ROUJ/) ROU J.
PA (KWAC/) KWACK S.
PA (CHOI/) CHOI I.
PA (KIMB/) KIM B.
XX
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;
XX
DR WPI; 2005-403340/41.
XX
XX Functional cDNA chip useful for screening and function analysis of growth
XX specific genes according to breeds and tissues of swine, comprises
PT substrate and probe comprising growth specific genes in muscle and fat
PT tissues of swine.
XX
XX Claim 2; SEQ ID NO 5; 8pp; English.
XX
XX The invention relates to a functional cDNA chip (I) for screening and
CC function analysis of growth specific genes according to breeds and
CC tissues of swine, comprises a probe comprising growth specific genes in
CC muscle and fat tissues of swine, and a substrate on which the probe is
CC immobilized. Further disclosed is a kit (KI) useful for screening and
CC functional analysis of growth specific gene according to breeds and
CC tissues of swine, comprising (I) integrated in it, Cys-dCTP or Cys-dCTP
CC bound cDNA from RNA of the tissue to be screened, a fluorescence scanning
CC system, and a computer analysis system. (I) and (KI) are useful for
CC screening and function analysis of growth specific gene according to
CC breeds and tissues of swine. (II) is useful in the swine improvement and
CC breeding of a new breed, and in the hog raising industry. The current
CC sequence represents the pig growth factor V (GF V) gene.
XX
XX Sequence 507 BP; 117 A; 160 C; 120 G; 110 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 507; DB 14; Length 507;
Best Local Similarity 100.0%; Pred. No. 9.1e-153;
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TATATAGAACCGAATCAGTACATCGGCTTGACCAAGCAGGCGCAAAACAGGCAACCT 60
Db 1 TATATAGAACCGAATCAGTACATCGGCTTGACCAAGCAGGCGCAAAACAGGCAACCT 60
Qy 61 AGGAGGTTTATAAATAGTATACCGCGCTGACACATACATACACTACTACCCGAAACCGG 120
Db 61 AGGAGGTTTATAAATAGTATACCGCGCTGACACATACATACACTACTACCCGAAACCGG 120
Qy 121 GGACAACTAGGGCTCCGCGCATAGCCATCCCTTTCTGTCGTCGATGTTGGGCTGCAG 180
Db 121 GGACAACTAGGGCTCCGCGCATAGCCATCCCTTTCTGTCGTCGATGTTGGGCTGCAG 180
Qy 181 TTATAGGCTGCGCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240
Db 181 TTATAGGCTGCGCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240
Qy 241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGCGCTGAGCGCTTACCCCTGCGCAG 300
Db 241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGCGCTGAGCGCTTACCCCTGCGCAG 300
Qy 301 TGTGGTGTAGCTCCGCTTAAAGCTTCCCGATAGCGCGCTTTTACACACCATCGG 360
Db 301 TGTGGTGTAGCTCCGCTTAAAGCTTCCCGATAGCGCGCTTTTACACACCATCGG 360
Qy 361 CGGACTAGACACCGTTGTTGAGCGTAAAGCGTCTATGGTAGAGCTGCGCGGACCGCG 420
Db 361 CGGACTAGACACCGTTGTTGAGCGTAAAGCGTCTATGGTAGAGCTGCGCGGACCGCG 420
Qy 421 TGTAGCAGCTTACTACATGTTAGTTTCAGCAACACCCCTCCCAATCCCGTGTCCCTA 480
Db 421 TGTAGCAGCTTACTACATGTTAGTTTCAGCAACACCCCTCCCAATCCCGTGTCCCTA 480
Qy 481 CTCCAACCTCTGTCGGTTTCAGCCGCGAG 507
Db 481 CTCCAACCTCTGTCGGTTTCAGCCGCGAG 507

```

---

```

RESULT 2
ADZ79335
ID ADZ79335 standard; cDNA; 507 BP.
XX
AC ADZ79335;
XX
DT 28-JUL-2005 (first entry)
XX
XX Swine growth factor GF-V cDNA for cDNA chip diagnostic method.
XX
XX biochip; screening; EST; expressed sequence tag; probe; muscle;
KW immobilization; gene expression; polymorphism; diagnosis;
KW animal breeding; growth factor; ss.
XX
XX Unidentified.
XX
XX US2005112602-A1.
XX
XX 26-MAY-2005.
XX
XX 27-FEB-2004; 2004US-00789723.
XX
XX 24-NOV-2003; 2003KR-00083651.
XX
XX (KIMC/) KIM C.
PA (YEOU/) YEO J.
PA (LEEJ/) LEE J.
PA (SONG/) SONG Y.
PA (CHOK/) CHO K.
PA (CHUN/) CHUNG K.
PA (KIMI/) KIM I.
PA (JINS/) JIN S.
PA (PARK/) PARK S.
PA (JUNG/) JUNG J.
PA (LEEM/) LEE M.
PA (KWON/) KWON E.
PA (CHOE/) CHO B.
PA (CHOH/) CHO H.
PA (SHIN/) SHIN S.
PA (NAMH/) NAM H.
PA (HONG/) HONG Y.
PA (KANG/) KANG Y.
PA (HAYY/) HA Y.
PA (ROUJ/) ROU J.
PA (KWAC/) KWACK S.
PA (CHOI/) CHOI I.
PA (KIMB/) KIM B.
XX
XX Kim C, Yeo J, Lee J, Song Y, Cho K, Chung K, Kim I, Jin S;
PI Park S, Jung J, Lee M, Kwon E, Cho E, Cho H, Shin S, Nam H;
PI Hong Y, Hong S, Kang Y, Ha Y, Rou J, Kwack S, Choi I, Kim B;
XX
XX WPI; 2005-371656/38.
XX
XX New cDNA chip comprising a probe capable of detecting marker genes
PT specifically expressed in the muscle and fat tissues of swine, and a
PT substrate, useful for screening and function analysis of swine genes.
XX
XX Claim 9; SEQ ID NO 5; 15pp; English.
XX
XX The invention relates to a cDNA chip for screening and function analysis
CC of swine genes comprising a probe capable of detecting marker genes
CC specifically expressed in the muscle and fat tissues of swine and a
CC substrate on which the probe is immobilized. The invention also includes
CC a kit for screening and function analysis of swine genes comprising the
CC cDNA chip. The cDNA chip is useful for screening and function analysis of
CC swine genes, particularly for comparing genetic expression according to
CC swine breeds and tissues, genetic mutation screening, genetic
CC polymorphism interpretation, development of new drugs for disease
CC treatment and disease diagnosis, swine improvement. This sequence

```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocolloration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 10:51:20 ; Search time 3003.91 Seconds  
(without alignments)  
7896.735 Million cell updates/sec

Title: US-10-785-981-5  
Perfect score: 507  
Sequence: 1 tatatagaaccgaatcacgt.....tctgtcggttcagcgcgag 507

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	12.0	932	5	BX432596
2	41.4	8.2	529	1	AV429458
3	41.2	8.1	670	7	CJ019297
4	41.2	8.1	701	3	BP169281
5	41.2	8.1	707	7	CJ016021
6	41.2	8.1	794	3	BP144350
7	41.2	8.1	803	8	DN106813
8	41.2	8.1	830	3	BP156807
9	41.2	8.1	862	7	CJ014151
10	41.2	8.1	883	8	DN107179
11	41.2	8.1	896	9	CC086053
12	40.8	8.0	1180	8	DN703457
13	40.2	7.9	649	1	AU296640
14	40.2	7.9	813	7	CO448190
15	39.6	7.8	644	5	BW964713
16	39.6	7.8	692	3	BP153784
17	39.4	7.8	689	6	CF076027
18	39.2	7.7	490	7	CO975370
19	39.2	7.7	536	7	CO974837
20	39.2	7.7	567	7	CO972208
21	39.2	7.7	578	7	CN868859
22	39.2	7.7	907	7	CO451557

C 23	39.2	7.7	1069	7	CK217376
C 24	39.2	7.7	1192	6	CD499604
C 25	39	7.7	572	8	CX685065
C 26	39	7.7	929	8	DN567210
C 27	38.6	7.6	412	6	CD489728
C 28	38.6	7.6	479	9	BH211305
C 29	38.6	7.6	518	9	AQ977896
C 30	38.6	7.6	704	7	CO452019
C 31	38.6	7.6	787	7	CO445849
C 32	38.6	7.6	803	7	CO445036
C 33	38.6	7.6	837	7	CO446224
C 34	38.6	7.6	853	7	CO463295
C 35	38.4	7.6	401	7	CN010492
C 36	38.4	7.6	448	7	CN010179
C 37	38.4	7.6	486	7	CF842501
C 38	38.4	7.6	494	7	CF841197
C 39	38.4	7.6	495	7	CF851149
C 40	38.4	7.6	593	7	CF839520
C 41	38.4	7.6	682	7	CF840809
C 42	38.4	7.6	688	7	CF864916
C 43	38.4	7.6	706	7	CN011313
C 44	38.4	7.6	715	7	CF851215
C 45	38.2	7.5	461	7	CV515593

## ALIGNMENTS

RESULT 1  
BX432596  
LOCUS  
DEFINITION BX432596 Homo sapiens FETAL BRAIN Homo sapiens cdna clone  
ACCESSION BX432596  
VERSION BX432596.1 GI:30779118  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 932)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
JOURNAL  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
This sequence belongs to sequence cluster 3874.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0BAH009ZB10NPI&c=3874.r.

## FEATURES

Location/Qualifiers  
1..932  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DP032YF14"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

CO451557 MZCCL1016



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:15:29 ; Search time 124.404 Seconds  
(without alignments)  
7244.353 Million cell updates/sec

Title: US-10-785-981-5

Perfect score: 507

Sequence: 1 tatatagaaccgaacacgt.....tctgtcggttcagccgcag 507

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PTUS COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.4	8.2	336	2	US-07-814-220-3
C 2	41.4	8.2	336	2	US-07-812-421-3
C 3	39.8	7.9	336	2	US-07-814-220-4
C 4	39.8	7.9	336	2	US-07-812-421-4
C 5	39	7.7	7218	2	US-08-232-463-14
C 6	36	7.1	4403765	3	US-09-103-840A-2
C 7	34.8	6.9	4411529	3	US-09-103-840A-1
C 8	34.6	6.8	336	2	US-07-814-220-3
C 9	34.6	6.8	336	2	US-07-812-421-3
C 10	34	6.7	1879	2	US-08-403-852D-5
C 11	34	6.7	1879	2	US-08-310-646B-5
C 12	34	6.7	1879	3	US-09-231-818-5
C 13	34	6.7	1879	3	US-09-635-359B-5
C 14	33.8	6.7	1294	3	US-09-312-038-4
C 15	33.8	6.7	1294	3	US-09-850-964-4
C 16	33.8	6.7	2289	3	US-09-312-038-3
C 17	33.8	6.7	2289	3	US-09-850-964-3
C 18	33.6	6.6	1505	3	US-09-270-767-11623
C 19	33.2	6.5	567	4	US-09-605-703B-1299
C 20	33	6.5	336	2	US-07-814-220-4
C 21	33	6.5	336	2	US-07-812-421-4
C 22	32.6	6.4	700	3	US-09-735-271-1138
C 23	32.4	6.4	2689	2	US-08-465-795-2
C 24	32.4	6.4	5589	2	US-08-465-795-1

## ALIGNMENTS

RESULT 1  
US-07-814-220-3/c  
; Sequence 3, Application US/07814220  
; Patent No. 5925540  
; GENERAL INFORMATION:  
; APPLICANT: Caceci, Thomas  
; APPLICANT: Toth, Thomas E.  
; APPLICANT: Szumanski, Maria B.W.  
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND  
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM  
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,  
; CITY: Reston  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 20191  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/814,220  
; FILING DATE: 23-DEC-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/588,437  
; FILING DATE: 25-SEP-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Whitham, Michael E.  
; REGISTRATION NUMBER: 32,635  
; REFERENCE/DOCKET NUMBER: CIT.016  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-391-2510  
; TELEFAX: 703-391-9035  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 336 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS

Sequence 5595, Ap  
Sequence 7538, Ap  
Sequence 2820, A  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 7512, Ap  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 5303, Ap  
Sequence 5536, Ap  
Sequence 6394, Ap  
Sequence 11950, A  
Sequence 12062, A  
Sequence 11981, A  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 8976, Ap  
Sequence 5780, Ap  
Sequence 890, App  
Sequence 165855,

354 3 US-09-248-796A-5595  
516 3 US-09-270-767-7538  
516 3 US-09-270-767-22820  
477 3 US-09-135-994-1  
477 3 US-09-584-843A-1  
672 3 US-09-248-796A-7512  
6530 2 US-08-146-930-1  
6530 3 US-08-458-240-1  
6530 6 PCT-US93-03993-1  
639 3 US-09-489-039A-5903  
861 3 US-09-489-039A-5536  
1410 3 US-09-489-039A-6394  
1215 3 US-09-252-991A-11950  
1365 3 US-09-252-991A-12062  
3390 3 US-09-252-991A-11981  
4403765 3 US-09-103-840A-2  
4411529 3 US-09-103-840A-1  
399 3 US-09-621-976-8976  
1059 3 US-09-489-039A-5780  
369 3 US-09-543-681A-890  
601 3 US-09-949-016-165855

```

; LOCATION: 31..324
US-07-814-220-3

Query Match      8.2%; Score 41.4; DB 2; Length 336;
Best Local Similarity 49.8%; Pred. No. 0.0007;
Matches 105; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 297 GCAGTGTGGTGTAGCTCCGCTCTAAGCTTCCCGATAGCCGCGCTTTTACACACCA 356
    |||||
Db 248 GCAGTGTGGTGTAGCTCCGCGAGCGCTGTGCGGTGCGCGAGTTGCTGCGCGAGCA 189
    |||||
QY 357 TCGCGGAGTACACACCGTTGGTTGCAGCGTAAAGCTCTATGTAGTAGCAGTGGCGGACC 416
    |||||
Db 188 GCTGCTGCGGTAGCAGCGGTAGCTGCGCGAGCGCGCTGCTGCGGTGCGCGAGTAGCTGCC 129
    |||||
QY 417 GCGGTGTAGCAGCTTACTATGTTAGTTTTCAGCAACCAACCCTTGCCCAATACCGGTGTT 476
    |||||
Db 128 GCAGCGCTGCTGCGGTGCGCGAGTAGCTGCGCGAGCGCGCTGCTGCGGTGCGCGAGTA 69
    |||||
QY 477 CCTACTCCAACCTCTGTCGGTTTCAGCGCGAG 507
    |||||
Db 68 GCAGCTGCTGCTGCCGCGCGTGGGAAGCG 38
    |||||
```

```

RESULT 2
US-07-812-421-3/c
; Sequence 3, Application US/07812421
; Patent No. 5932697
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; CITY: Suite 900
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/812,421
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..324
; US-07-812-421-3

Query Match      7.9%; Score 39.8; DB 2; Length 336;
Best Local Similarity 49.3%; Pred. No. 0.0025;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 297 GCAGTGTGGTGTAGCTCCGCTCTAAGCTTCCCGATAGCCGCGCTTTTACACACCA 356
    |||||
Db 248 GCAGTGTGGTGTAGCTCCGCGAGCGCTGTGCGGTGCGCGAGTTGCTGCGCGAGCA 189
    |||||
QY 357 TCGCGGAGTACACACCGTTGGTTGCAGCGTAAAGCTCTATGTAGTAGCAGTGGCGGACC 416
    |||||
Db 188 GCTGCTGCGGTAGCAGCGGTAGCTGCGCGAGCGCGCTGCTGCGGTGCGCGAGTAGCTGCC 129
    |||||
QY 417 GCGGTGTAGCAGCTTACTATGTTAGTTTTCAGCAACCAACCCTTGCCCAATACCGGTGTT 476
    |||||
Db 128 GCAGCGCTGCTGCGGTGCGCGAGTAGCTGCGCGAGCGCGCTGCTGCGGTGCGCGAGTA 69
    |||||
QY 477 CCTACTCCAACCTCTGTCGGTTTCAGCGCGAG 507
    |||||
Db 68 GCAGCTGCTGCTGCCGCGCGTGGGAAGCG 38
    |||||
```

```

RESULT 3
US-07-814-220-4
; Sequence 4, Application US/07814220
; Patent No. 5925540
; GENERAL INFORMATION:
; APPLICANT: Caceci, Thomas
; APPLICANT: Toth, Thomas E.
; APPLICANT: Szumanski, Maria B.W.
; TITLE OF INVENTION: SYNTHETIC ANTIFREEZE PEPTIDE AND
; TITLE OF INVENTION: SYNTHETIC GENE CODING FOR ITS PRODUCTION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WHITHAM, CURTIS & WHITHAM
; STREET: Reston Intl. Center, 11800 Sunrise Valley Dr.,
; CITY: Suite 900
; STATE: VA
; COUNTRY: USA
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/814,220
; FILING DATE: 23-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/588,437
; FILING DATE: 25-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitham, Michael E.
; REGISTRATION NUMBER: 32,635
; REFERENCE/DOCKET NUMBER: CIT.016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-391-2510
; TELEFAX: 703-391-9035
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 336 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-07-814-220-4

Query Match      7.9%; Score 39.8; DB 2; Length 336;
Best Local Similarity 49.3%; Pred. No. 0.0025;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

US-07-812-421-3
```

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:17:34 ; Search time 613.737 Seconds  
(without alignments)  
6831.225 Million cell updates/sec

Title: US-10-785-981-5

Perfect score: 507  
Sequence: 1 tatatagaacggaatcacgt.....tctgtcggttcagccgcag 507

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	507	100.0	507	9	US-10-785-981-5
2	507	100.0	507	9	US-10-789-723-5
3	507	100.0	507	9	US-10-786-052-5
4	123.2	24.3	505	7	US-10-424-599-88182
5	61	12.0	523	7	US-10-424-599-79262
6	48.4	9.5	520	5	US-10-184-644-332
7	48.4	9.5	520	5	US-10-184-634-332
8	39.8	7.9	2852	10	US-11-097-143-1178
9	39.8	7.9	5023	10	US-11-097-143-1177
10	39.6	7.8	5093	7	US-10-424-599-134445
11	39.6	7.8	5643	6	US-10-012-952A-43
12	38.6	7.6	3234	8	US-10-425-115-172916
13	38	7.5	480	8	US-10-425-115-167165
14	38	7.5	675	7	US-10-437-963-1319
15	37.2	7.3	2712	7	US-10-282-122A-24663
16	36.6	7.2	610	7	US-10-437-963-68546
17	36.6	7.2	1888	7	US-10-437-963-73429
18	36	7.1	728	7	US-10-398-221-1694
19	36	7.1	1296	7	US-10-398-221-610
20	36	7.1	1311	7	US-10-398-221-613
21	36	7.1	2034	7	US-10-398-221-3554
22	36	7.1	319630	7	US-10-398-221-7
23	36	7.1	3011208	7	US-10-398-221-2058

C	24	35.8	7.1	653	5	US-10-184-644-402	Sequence 402, App
C	25	35.8	7.1	653	5	US-10-184-634-402	Sequence 402, App
C	26	35.8	7.1	1933	8	US-10-425-115-29072	Sequence 29072, A
C	27	35.4	7.0	539	7	US-10-425-114-5025	Sequence 5025, Ap
C	28	35.4	7.0	561	7	US-10-425-114-17347	Sequence 17347, A
C	29	35.4	7.0	642	7	US-10-425-114-8863	Sequence 8863, Ap
C	30	35.4	7.0	856	6	US-10-310-154-35	Sequence 35, Appl
C	31	35	6.9	863	3	US-09-974-300-5249	Sequence 5249, Ap
C	32	35	6.9	2511	9	US-10-846-374B-51	Sequence 51, Appl
C	33	35	6.9	9732	9	US-10-450-763-15494	Sequence 15494, A
C	34	34.8	6.9	218802	9	US-10-897-508-1	Sequence 1, Appli
C	35	34.6	6.8	408	7	US-10-282-122A-26209	Sequence 26209, A
C	36	34.6	6.8	411	7	US-10-282-122A-28249	Sequence 28249, A
C	37	34.6	6.8	1010	9	US-10-450-763-4066	Sequence 4066, Ap
C	38	34.6	6.8	1167	7	US-10-282-122A-42006	Sequence 42006, A
C	39	34.6	6.8	1599	7	US-10-437-963-57512	Sequence 57512, A
C	40	34.6	6.8	2466	7	US-10-468-250A-74	Sequence 74, Appl
C	41	34.4	6.8	824	3	US-09-966-881-54	Sequence 54, Appl
C	42	34.4	6.8	1062	7	US-10-282-122A-13433	Sequence 13433, A
C	43	34.2	6.7	718	8	US-10-425-115-44833	Sequence 44833, A
C	44	34.2	6.7	752	7	US-10-767-701-8355	Sequence 8355, Ap
C	45	34	6.7	1449	3	US-09-738-626-1597	Sequence 1597, Ap

ALIGNMENTS

RESULT 1  
US-10-785-981-5  
; Sequence 5, Application US/10785981  
; Publication No. US20050112597A1  
; GENERAL INFORMATION:  
; APPLICANT: GYRONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: SCREENING EXPRESSION PROFILE OF GROWTH SPECIFIC GENES IN SWINE  
; TITLE OF INVENTION: AND FUNCTIONAL CDNA CHIP PREPARED BY USING THE SAME  
; FILE REFERENCE: 3884-0120P  
; CURRENT FILING DATE: 2004-06-26  
; CURRENT FILING DATE: 2004-06-26  
; PRIOR APPLICATION NUMBER: KR 2003-83653  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 5  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Kegoshima Berkshire  
US-10-785-981-5

Query Match	100.0%	Score 507;	DB 9;	Length 507;
Best Local Similarity	100.0%	Pred. No. 2.5e-161;		
Matches 507;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TATATAGAACCGAATCAGTACCTGGGCTGTACCAAGCAGGCGCCAAACAAAGCAACCT	60	
Db	1	TATATAGAACCGAATCAGTACCTGGGCTGTACCAAGCAGGCGCCAAACAAAGCAACCT	60	
Qy	61	AGGAGGTATATAAATAGGTATACGCGCTGTACACATACATACCTACTACCCGAACCGG	120	
Db	61	AGGAGGTATATAAATAGGTATACGCGCTGTACACATACATACCTACTACCCGAACCGG	120	
Qy	121	GGACAACTAGGGCTCCGCCATAGCCATCTTTCTTGGTGTGCATGTTTGGGGGTGCAG	180	
Db	121	GGACAACTAGGGCTCCGCCATAGCCATCTTTCTTGGTGTGCATGTTTGGGGGTGCAG	180	
Qy	181	TTATAGGGCTGCCAACCGCCATACACCTTACAGGCACCTTATTAGTTACATCCACGA	240	
Db	181	TTATAGGGCTGCCAACCGCCATACACCTTACAGGCACCTTATTAGTTACATCCACGA	240	
Qy	241	GGGCTCTGTACCAACCCCTAAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG	300	
Db	241	GGGCTCTGTACCAACCCCTAAGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAGTGGCAG	300	

Qy 301 TGTGGTGTAGCTCCGTCCTAAGCTTCCCGATAGCGCGCGTCTTTTACACACCATCGG 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 TGTGGTGTAGCTCCGTCCTAAGCTTCCCGATAGCGCGCGTCTTTTACACACCATCGG 360  
Qy 361 CGGACTAGACACCGCTTGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 CGGACTAGACACCGCTTGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Qy 421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Qy 481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507

RESULT 2  
US-10-789-723-5  
; Sequence 5, Application US/10789723  
; Publication No. US20050112602A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: cDNA chip for screening specific genes and analyzing their  
; TITLE OF INVENTION: function in swine  
; FILE REFERENCE: YLOP040109US  
; CURRENT APPLICATION NUMBER: US/10789, 723  
; CURRENT FILING DATE: 2004-02-27  
; PRIOR APPLICATION NUMBER: KR 2003-83651  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 5  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-789-723-5

Query Match 100.0%; Score 507; DB 9; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2.5e-161;  
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATATAGAACCGAATCAGCTACACTGGGCTTGACCAAGCAGGCGCCAAACAAAGCAACCT 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 TATATAGAACCGAATCAGCTACACTGGGCTTGACCAAGCAGGCGCCAAACAAAGCAACCT 60  
Qy 61 AGGAGGTTTATAAATAGGTATACGCGCGCTGACACATACATCTACTACCCGAAACGCGG 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 AGGAGGTTTATAAATAGGTATACGCGCGCTGACACATACATCTACTACCCGAAACGCGG 120  
Qy 121 GGACAACTAGGGCTCCGCCATTAAGCCATCCCTTCTCGTGCATGTTGGGCTGCAG 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 GGACAACTAGGGCTCCGCCATTAAGCCATCCCTTCTCGTGCATGTTGGGCTGCAG 180  
Qy 181 TTATAGGCTGCCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 TTATAGGCTGCCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240  
Qy 241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGTAGCGGCTTACCCCTGCGCAG 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGTAGCGGCTTACCCCTGCGCAG 300  
Qy 301 TGTGGTGTAGCTCCGTCCTAAGCCTTCCCGATAGCGCGCTTTTACACACATCGG 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 TGTGGTGTAGCTCCGTCCTAAGCCTTCCCGATAGCGCGCTTTTACACACATCGG 360  
Qy 361 CGGACTAGACACCGTTCGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 CGGACTAGACACCGTTCGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Qy 421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480

Db 421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Qy 481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507

RESULT 3  
US-10-786-052-5  
; Sequence 5, Application US/10786052  
; Publication No. US20050113568A1  
; GENERAL INFORMATION:  
; APPLICANT: GYEONGSANGNAM-DO  
; APPLICANT: Chulwook, KIM  
; TITLE OF INVENTION: NOVEL GROWTH RELATED GENES FROM SWINE  
; FILE REFERENCE: 3884-0119P  
; CURRENT APPLICATION NUMBER: US/10786, 052  
; CURRENT FILING DATE: 2004-02-26  
; PRIOR APPLICATION NUMBER: KR 2003-83652  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 5  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Kagoshima Berkshire  
US-10-786-052-5

Query Match 100.0%; Score 507; DB 9; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2.5e-161;  
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATATAGAACCGAATCAGCTACACTGGGCTTGACCAAGCAGGCGCCAAACAAAGCAACCT 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 TATATAGAACCGAATCAGCTACACTGGGCTTGACCAAGCAGGCGCCAAACAAAGCAACCT 60  
Qy 61 AGGAGGTTTATAAATAGGTATACGCGCGCTGACACATACATCTACTACCCGAAACGCGG 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 AGGAGGTTTATAAATAGGTATACGCGCGCTGACACATACATCTACTACCCGAAACGCGG 120  
Qy 121 GGACAACTAGGGCTCCGCCATTAAGCCATCCCTTCTCGTGCATGTTGGGCTGCAG 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 GGACAACTAGGGCTCCGCCATTAAGCCATCCCTTCTCGTGCATGTTGGGCTGCAG 180  
Qy 181 TTATAGGCTGCCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
181 TTATAGGCTGCCAACCGCCATACACACTTACAGCCACTTATTAAGTTACATCCACGA 240  
Qy 241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGTAGCGGCTTACCCCTGCGCAG 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 GGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGTAGCGGCTTACCCCTGCGCAG 300  
Qy 301 TGTGGTGTAGCTCCGTCCTAAGCCTTCCCGATAGCGCGCTTTTACACACATCGG 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 TGTGGTGTAGCTCCGTCCTAAGCCTTCCCGATAGCGCGCTTTTACACACATCGG 360  
Qy 361 CGGACTAGACACCGTTCGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 CGGACTAGACACCGTTCGGTTGCGAGCGTAAGCGTCTATGGTAGCAGCTGCGGCGACCGCG 420  
Qy 421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 TGAGCCAGCTTACTACATGTTAGTTTCAGCAACCCCTGCGCAATACCCGTGTTCCCTA 480  
Qy 481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 CTCCAACTCTGTCGGTTTCAGCCGCGAG 507

RESULT 4  
US-10-424-599-88182  
; Sequence 88182, Application US/10424599  
; Publication No. US20040031072A1



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 20, 2006, 11:27:30 ; Search time 387.381 Seconds  
(without alignments)  
2778.314 Million cell updates/sec

Title: US-10-785-981-5

Perfect score: 507

Sequence: 1 tatatagaaccgaatcacgt.....ctctgctgttcagccgcag 507

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204323 seqs, 1061406715 residues

Total number of hits satisfying chosen parameters: 14408646

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA New.\*
- 1: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.8	6.9	3990	12	US-11-052-554A-520 Sequence 520, App
2	32.2	6.4	2539	12	US-11-136-527-2523 Sequence 2523, App
3	32	6.3	523	6	US-09-925-065A-240456 Sequence 240456, App
4	32	6.3	617	6	US-09-925-065A-461295 Sequence 461295, App
5	32	6.3	993	12	US-11-128-061-930 Sequence 930, App
6	32	6.3	993	12	US-11-128-049-930 Sequence 930, App
c 7	32	6.3	1512	6	US-09-925-065A-62278 Sequence 62278, A
c 8	32	6.3	2223	8	US-10-467-657-6265 Sequence 6265, App
c 9	32	6.3	2539	12	US-11-136-527-215 Sequence 215, App
c 10	31.8	6.3	2479	12	US-11-136-527-3303 Sequence 3303, App
11	31.4	6.2	476	6	US-09-925-065A-472363 Sequence 472363, App
12	31.4	6.2	1090	6	US-09-925-065A-680565 Sequence 680565, App
c 13	31.2	6.2	619	6	US-09-925-065A-874582 Sequence 874582, App
c 14	31.2	6.2	978	8	US-10-793-626-3295 Sequence 3295, App
c 15	31.2	6.2	2993	8	US-10-793-626-4155 Sequence 4155, App
c 16	31.2	6.2	3585	8	US-10-793-626-3521 Sequence 3521, App
c 17	31	6.1	255	8	US-10-909-125-1963 Sequence 1963, App
c 18	31	6.1	889	12	US-11-136-527-373 Sequence 373, App
c 19	31	6.1	889	12	US-11-136-527-4469 Sequence 4469, App
c 20	31	6.1	2183	9	US-11-072-512-1064 Sequence 1064, App

ALIGNMENTS

RESULT 1

US-11-052-554A-520  
; Sequence 520, Application US/11052554A  
; Publication No. US20050288866A1  
; GENERAL INFORMATION:  
; APPLICANT: Sachdeva, et al.  
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
; FILE REFERENCE: 30853/40359A  
; CURRENT APPLICATION NUMBER: US/11/052,554A  
; CURRENT FILING DATE: 2005-02-07  
; PRIOR APPLICATION NUMBER: US 60/589,227  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
; PRIOR FILING DATE: 2004-02-06  
; NUMBER OF SEQ ID NOS: 763  
; SOFTWARE: Patentin version 3.3  
; SEQ ID NO 520  
; LENGTH: 3990  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis H37Rv  
US-11-052-554A-520

Query Match 6.9%; Score 34.8; DB 12; Length 3990;  
Best Local Similarity 43.1%; Pred. No. 0.79;  
Matches 168; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

Qy	118	CGGGGCACTAGGGCTCGGCCATAGCCATCTTTCTGTCGTCGATGTGGGCTG	177
Db	3048	CGAGGCAAGCGGGCGCGCGGAGCATCTTGCACGCGGGCAACGGCGTACGG	3107
Qy	178	CAGTTATAGGGCTGCCAACCGCCATACACACTTTACCAGCCACTTATTAAAGTTACATCCA	237
Db	3108	CGGCCACGGCGCTCCAGCGCGGAAACGGCGCGCCCGCGGCAACGGCGTACGGCGG	3167
Qy	238	CGAGGGCTCTGTACACCCCTTAAGCAGTGGCAGTGGTAGCGCGTCCCGCTTACCTTCGG	297
Db	3168	TACGGCGGCAACGGCGCGGCAACGGCGGCAACGGCGGCAACGGCGGCAACGGCGGCGG	3227
Qy	298	CAGTGTGTGTCAGTCTCGTCTTCCCGATAGCTTCCCGATAGCGCGCGCTTTTACACACAT	357
Db	3228	CAGACCGGGCGCAAGGGCGTCTCTCGGCACCAAGGGCGATGGCGGTACGGCGCGAG	3287

```
Qy 358 CGCGGACTAGACACCGTTGGTTCGACGTAAGCGTCTATGGTAGCAGCTCGCGGACCG 417
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3288 AGCGGCAACGGCGGACAGCGGCAACGGCGGCAAGCGCTCACCGGAGCGCGGCAA 3347
Qy 418 CCGTGTAGCCAGCTTACTACATGTTAGTTTTCAGCAACCAACCCCTGCCAATACCCGTGTTC 477
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3348 CGGTGGGACCGGGGACACCGCGCAACGGCGGCAACGGTGGCAACGGCGCCAGTGGCGA 3407
Qy 478 CTACTCCACTCTGTCGGTTTCAGCCGAG 507
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3408 CCTTGTCACTCTACCTGGTGGACGGCGGG 3437

RESULT 2
US-11-136-527-2523
; Sequence 2523, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2523
; LENGTH: 2539
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-2523

Query Match 6.4%; Score 32.2; DB 12; Length 2539;
Best Local Similarity 48.6%; Pred. No. 4.7;
Matches 85; Conservative 1; Mismatches 89; Indels 0; Gaps 0;

Qy 326 TTCCCCGATAGCCCGCGCTTTTACACACCATCGCGGACGTAGACACCGTTGGTTGAGC 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
431 TCCCCGACGACCTCCCGGACCTTGTCTGCGAGCCTCAGCCTCCGAGCTTCAGCAGCAGC 490
Qy 386 GTAAGCGTCTATGTAGCAGCTCGCGGACCGCGGTGTAGCCAGCTTACTACATGTTAGT 445
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
491 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 550
Qy 446 TTCAGCAACCAACCTCGCAATACCCGTGTTCCCTACTCCCACTCTGTCGGTTTCA 500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
551 CCCTGCCTCAGCTTGCCCACTCCAGAGCCAGSTTGTCCTCCTGCTGTTGCA 605
```

```
RESULT 3
US-09-925-065A-240456
; Sequence 240456, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 240456
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-240456

Query Match 6.3%; Score 32; DB 6; Length 523;
Best Local Similarity 48.9%; Pred. No. 3.8;
Matches 86; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 44 CCAAAACAAGGCAACCTTAGGAGTTATAAAATAGGTATACGCGCGTGCACACATACATAC 103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
286 CAATCCATGCTGAGCTGAGAGTTGAAGGCTGTGCCCTATGCTGGCAGTCTCAAATAAAG 345
Qy 104 TCCTATCCCAACGCGGGGACAACTAGGCGTCCGCCATAGCCATCTTCTCTGGTGTGTC 163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
346 CAGCTAAGAGAGAGAGAAACCTTAGGCTCCAGCAGAGAGCCCTTCTGTCTCTGATATA 405
Qy 164 GATGTTGCGGCTGCAGTTATAGGCTGCCAACCGCCATACACACCTTACCAGCCA 219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
406 CCAGCTCAGGATAACTCTGTTAAGAGTCAATTTTGTGAACAAATCTGCAGAGCCA 461

RESULT 4
US-09-925-065A-461295
; Sequence 461295, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 461295
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-461295

Query Match 6.3%; Score 32; DB 6; Length 617;
Best Local Similarity 48.9%; Pred. No. 3.9;
Matches 86; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 44 CCAAAACAAGGCAACCTTAGGAGTTATAAAATAGGTATACGCGCGTGCACACATACATAC 103
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
10 CAATCCATGCTGAGCTGAGAGTTGAAGGCTGTGCCCTATGCTGGCAGTCTCAAATAAAG 69
Qy 104 TCCTATCCCAACGCGGGGACAACTAGGCTCCGCCATAGCCATCTTCTCTGGTGTGTC 163
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
70 CAGCTAAGAGAGAGAGAAACCTTAGGCTCCAGCAGAGAGCCCTTCTGTCTCTGATATA 129
Qy 164 GATGTTGCGGCTGCAGTTATAGGCTGCCAACCGCCATACACACCTTACCAGCCA 219
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 CCAGCTCAGGATAACTCTGTTAAGAGTCAATTTTGTGAACAAATCTGCAGAGCCA 185

RESULT 5
US-11-128-061-930
; Sequence 930, Application US/11128061
```